EXHIBIT 3

Exhibit 3

U.S. Patent No. 11,019,372 – Cisco Systems, Inc.

Claims 1, 6, 11

Scale Video Coding LLC ("SVC") provides evidence of infringement of claims 1, 6, and 11 of U.S. Patent No. 11,019,372 (hereinafter "the '372 patent") by Cisco Systems, Inc. ("Cisco"). In support thereof, SVC provides the following claim charts.

"Accused Instrumentalities" as used herein refers to at least the following Cisco products and services that implement the scalable video coding features of the H.264 video coding standard: Cisco Unified Communications Manager ("CUCM"); and Cisco Webex (including Cisco Webex Cloud, Webex Teams, Webex Meetings, Webex App, Webex Meetings Server, Webex Meeting Center, Webex Desk Pro, Webex Codec Pro, Webex Codec Plus, Webex WebRTC client, Webex Control Hub, Webex Video Mesh, Webex Cloud-Connected UC); along with associated hardware and/or software, including but not limited to other Cisco servers and related computer systems operated by Cisco that work in conjunction with CUCM and Cisco Webex.

These claim charts demonstrate Cisco's infringement by comparing each element of the asserted claims to corresponding components, aspects, and/or features of the Accused Instrumentalities. These claim charts are not intended to constitute an expert report on infringement. These claim charts include information provided by way of example, and not by way of limitation.

The analysis set forth below is based only upon information from publicly available resources regarding the Accused Instrumentalities, as Cisco has not yet provided any non-public information. An analysis of Cisco's (or other third parties') technical documentation and/or software source code may assist in fully identify all infringing features and functionality. Accordingly, SVC reserves the right to supplement this infringement analysis once such information is made available to SVC. Furthermore, SVC reserves the right to revise this infringement analysis, as appropriate, upon issuance of a court order construing any terms recited in the asserted claims.

SVC provides this evidence of infringement and related analysis without the benefit of claim construction or expert reports or the completion of fact discovery. SVC reserves the right to supplement, amend or otherwise modify this analysis and/or evidence based on any such claim construction or expert reports or additional discovery.

Unless otherwise noted, SVC contends that Cisco directly infringes the '372 patent in violation of 35 U.S.C. § 271(a) by selling, offering to sell, making, and/or using, the Accused Instrumentalities. The following exemplary analysis demonstrates that infringement. Cisco makes, uses, sells, imports, or offers for sale in the United States, or has made, used, sold, or offered for sale in the past, without

authority, without authority products, equipment, or services that infringe claims 1, 6, and 11 of the ’372 patent, including without limitation, the Accused Instrumentalities.

Unless otherwise noted, SVC believes and contends that each element of each claim asserted herein is literally met through Cisco’s provision of the Accused Instrumentalities. However, to the extent that Cisco attempts to allege that any asserted claim element is not literally met, SVC believes and contends that such elements are met under the doctrine of equivalents. More specifically, in its investigation and analysis of the Accused Instrumentalities, SVC did not identify any substantial differences between the elements of the patent claims and the corresponding features of the Accused Instrumentalities, as set forth herein. In each instance, the identified feature of the Accused Instrumentalities performs at least substantially the same function in substantially the same way to achieve substantially the same result as the corresponding claim element.

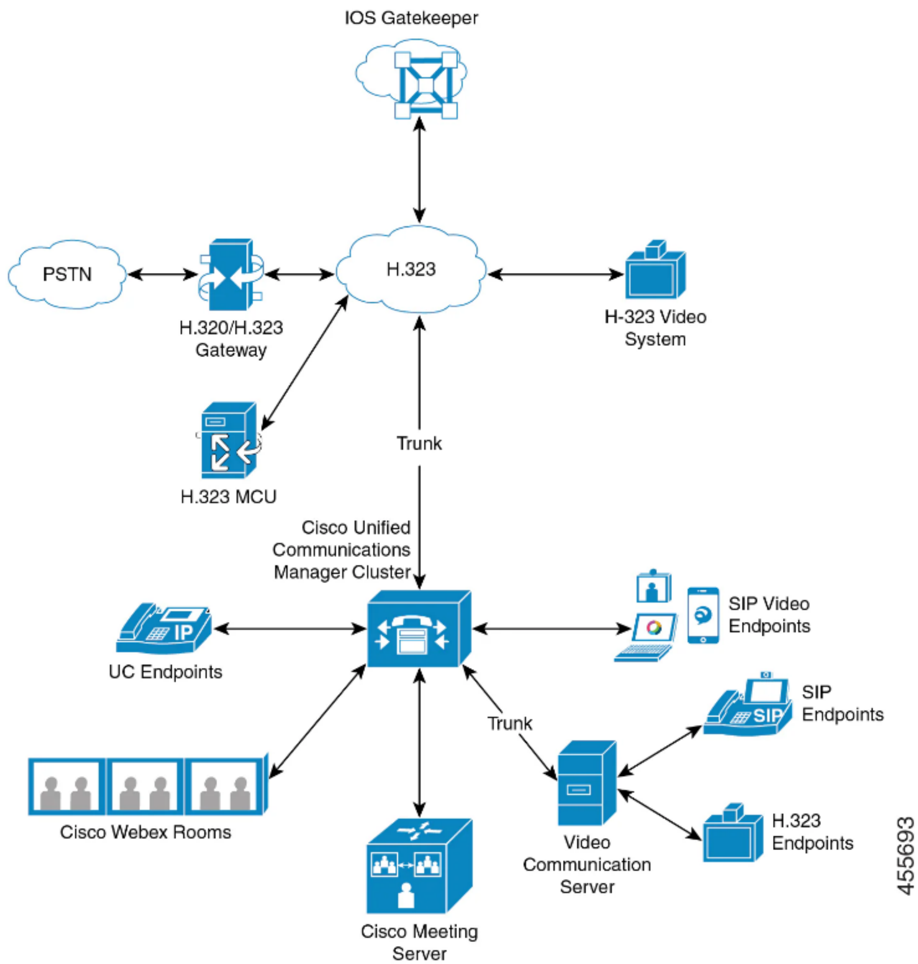
To the extent the chart of an asserted claim relies on evidence about certain specifically-identified Accused Instrumentalities, SVC asserts that, on information and belief, any similarly-functioning instrumentalities also infringe the charted claim. SVC reserves the right to amend this infringement analysis based on other products made, used, sold, imported, or offered for sale by Cisco. SVC also reserves the right to amend this infringement analysis by citing other claims of the ’372 patent, not listed in the claim chart, that are infringed by the Accused Instrumentalities. SVC further reserves the right to amend this infringement analysis by adding, subtracting, or otherwise modifying content in the “Accused Instrumentalities” column of each chart.

Claim Chart – Cisco

’372 Patent Claims 1, 6, 11	Accused Instrumentalities
1. A video router, comprising:	The Accused Instrumentalities include a video router. For example, scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding IP data packets containing scalable video information.

Video Network

The following illustration provides an example of a video network that uses a single Unified Communications Manager cluster. In a successful video network, any endpoint can call any other endpoint. Video availability only exists if both endpoints are video-enabled. Video capabilities extend across trunks.



	The Cisco video conference portfolio comprises the following video bridges: • Cisco TelePresence MCU series • Webex Meeting Server (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Figure 11-14 Cisco WebEx Collaboration Cloud Architecture

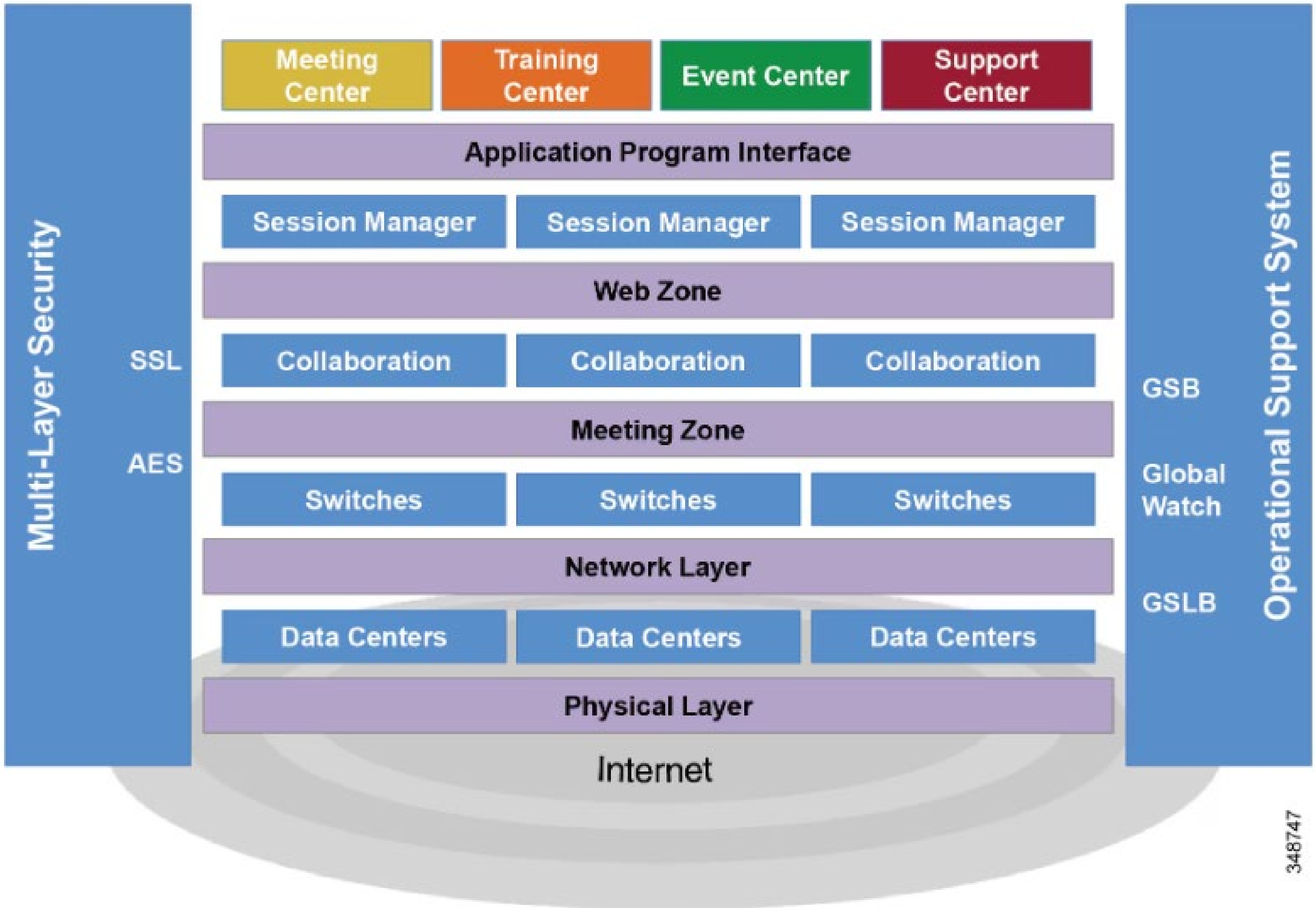
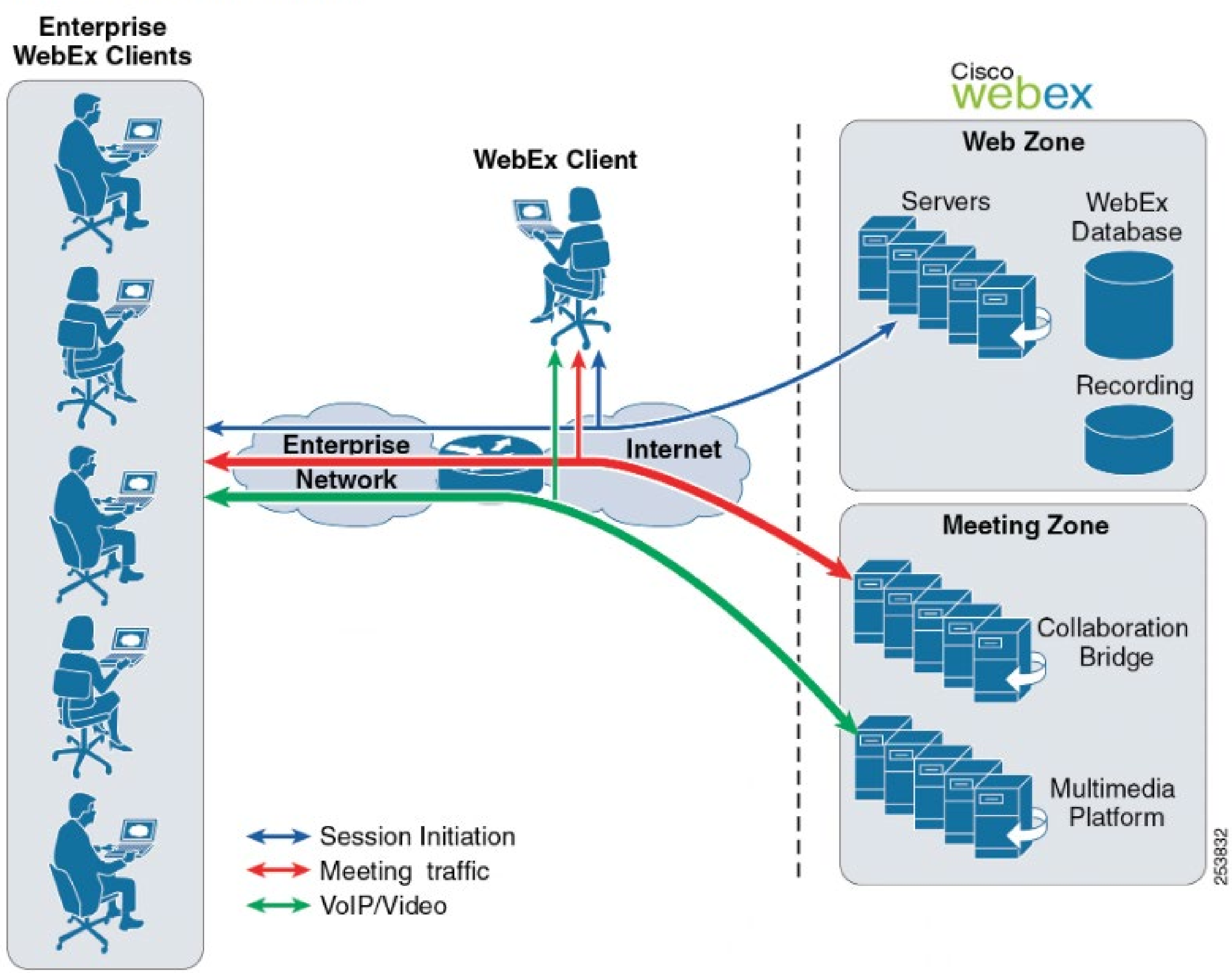


Figure 11-15 WebEx Deployment



	Cisco Rich Media Conferencing consists of the conferencing solutions described below. The details pertaining to each solution are described in each individual section that follows. • Cisco Unified CM Audio Conferencing This solution allows Unified CM to use its internal software component or external hardware digital signal processors (DSPs) as the resources to perform audio conferencing. • Cisco Meeting Server Cisco Meeting Server is an on-premises video conferencing solution. Each user has a personal Space that can be used to conduct meetings. Users can manage items such Space creation, adding members to a Space, and PIN creation from the Cisco Meeting App. • Cisco Collaboration Meeting Rooms Hybrid Cisco CMR Hybrid combines the on-premises video conference and the WebEx Meeting Center conference into a single meeting, which allows TelePresence and WebEx participants to join and share voice, video, and content. CMR Hybrid meetings can be either scheduled or non-scheduled. • Cisco WebEx Meeting Center Video Conferencing Cisco WebEx Meeting Center Video Conferencing (formerly Cisco Collaboration Meeting Rooms (CMR) Cloud) is an alternate conferencing deployment model that does not require any on-premises conferencing resources or management infrastructure. It supports both scheduled and non-scheduled meetings as well as TelePresence, audio, and WebEx participants in a single call, all hosted in the cloud. • Cisco WebEx Meetings Server Where cloud-based web and audio conferencing is not suitable, it is possible to use the on-premises WebEx Meetings Server solution. This product offers a standalone audio, video, and collaboration web conferencing platform. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)

	The Accused Instrumentalities implement the scalable video coding features in Annex G of the H.264 standard. For example: **Video Calls** The typical video call includes two or three Real-Time Protocol (RTP) streams in each direction (that is, four or six streams). The call can include the following stream types: • Video (H.261, H.263, H.263+, H.264-SVC, X-H.264UC, H.264-AVC, H.265, AV1 and VT Camera wideband video codecs) • Far-End Camera Control (FECC) - Optional • Binary Floor Control Protocols (BFCP) **SIP Video** SIP video supports the following video calls by using the SIP Signaling Interface (SSI): • SIP to SIP • SIP to H.323 • SIP to SCCP • SIP intercluster trunk • H.323 trunk • Combination of SIP and H.323 trunk SIP video calls also provide media control functions for video conferencing. Unified Communications Manager video supports SIP on both SIP trunks and lines support video signaling. SIP supports the H.261, H.263, H.263+, H.264 (AVC), H.264 (SVC), X-H.264UC (Lync), and AV1 video codecs (it does not support the wideband video codec that the VTA uses).

Video Codecs

Common video codecs include H.261, an older video codec, H.263, a newer codec that gets used to provide internet protocol (IP) video, and H.264, a high-quality codec. The system supports H.264 for calls that use the Skinny Client Control Protocol (SCCP), H.323, and SIP on originating and terminating endpoints only. The system also supports regions and locations.

Unified Communications Manager maintains the offerer's video codec ordering preference when making the answer, if possible. H.265 is the preferred video codec is available on the endpoints, otherwise, Unified Communications Manager follows the following codec preference preference order:

Preference Order	Codecs	Description
1	H.265 (HEVC)	Provides higher quality video using lower bandwidth.
2	H.264 (SVC)	Allows rendering of variable quality video from the same media stream, by disregarding a subset of the packets received. **Note** H. 264 SVC is a new annex to the H.264-AVC video compression standard; meaning it is an enhancement on top of H.264-AVC. It provides the ability to encapsulate multiple video streams at various frame-rates and resolutions in one container.

(https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

	Meeting Center uses the H.264 AVC/SVC codec to provide high-definition video for the conference. Higher network bandwidth is needed for those deployments. For further details regarding network traffic optimization for high-definition video, see Capacity Planning. **Network Traffic Planning** Network traffic planning for Cisco WebEx Meeting Center Video Conferencing consists of the following elements: • WebEx Clients bandwidth The WebEx meeting client uses Scalable Video Coding (SVC) technology to send and receive video. It uses multi-layer frames to send video, and the receiving client automatically selects the best possible resolution to receive video that typically requires 1.2 to 3 Mbps of available bandwidth. For more information regarding network traffic planning for WebEx clients, refer to the *Cisco WebEx Network Bandwidth* white paper, available at https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meeting-center/white_paper_c11-691351.html • Bandwidth for video device from enterprise to WebEx Cloud For optimal SIP audio and video quality, Cisco recommends setting up the video bandwidth for at least 1.5 Mbps per device screen in the region associated with the endpoint registering with Cisco Unified CM. For example, if a triple-screen device registers with Unified CM, video bandwidth of 4.5 Mbps should be allocated in the associated region. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)
a memory; and a processor, wherein the processor	The Accused Instrumentalities include a video router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to perform the claimed steps.

executes instructions stored in the memory to cause the video router to:	For example, all routers comprise a memory and a processor, wherein the processor executes instructions stored in the memory:

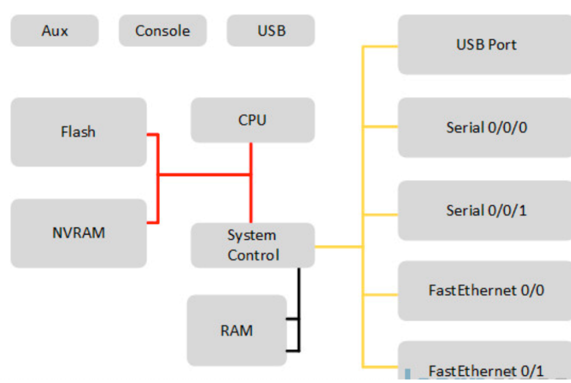
https://www.learncisco.net/courses/icnd-2/network-management/router-internal-components-overview.html

Router Internal Components Overview

I want you to take a moment and look at the components we have listed here.

- CPU - Executes operating system instructions
- RAM (Random access memory) - Contains the running copy of the configuration file. Stores the rougint table. RAM contents lost when power is off.
- ROM (Read-only memory) - Holds diagnostic software used when router is powered up. Stores the router's bootstrap program.
- NVRAM (Nonvolatile RAM) - Stores startup configuration. This may include IP addresses (routing protocol, hostname of a router).
- Flash memory - Contains the operating system (Cisco IOS)
- Interfaces - Multiple fisical interfaces that are used to connect the network. Fast Ethernet, Gigabit Ethernet; Serial; Management interfaces

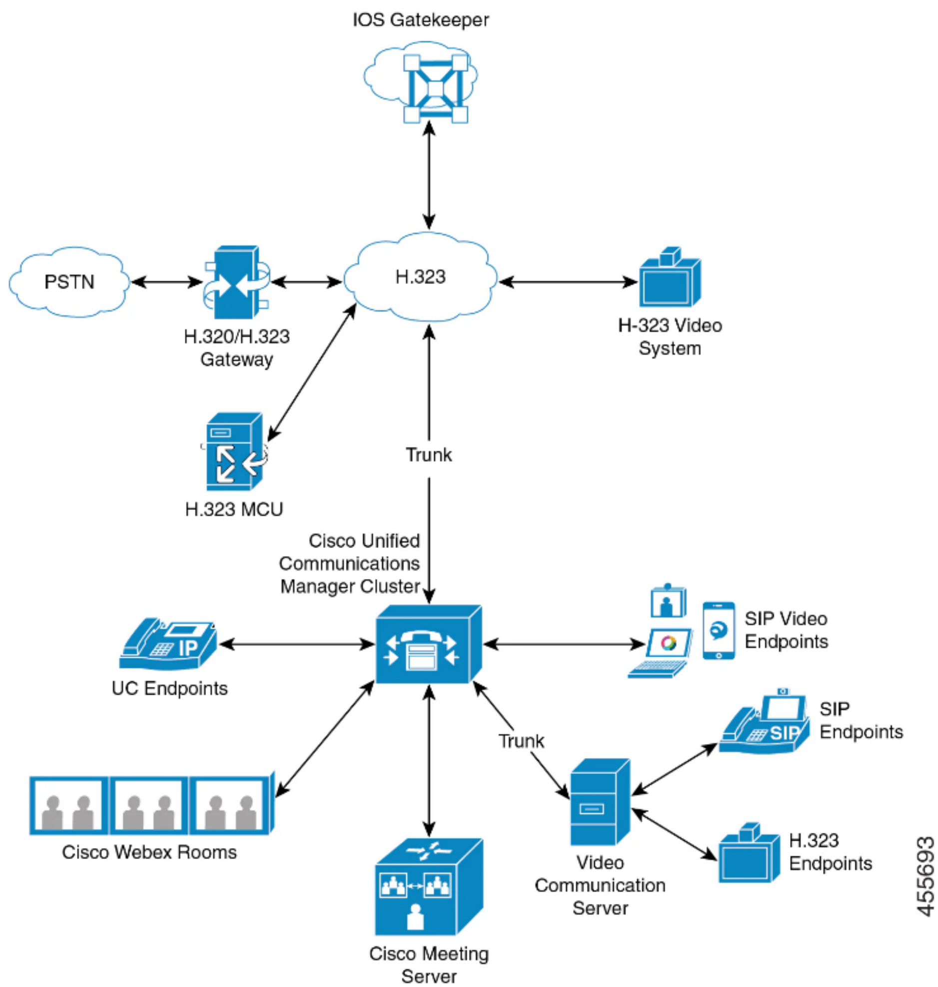
CPU, RAM, ROM, interfaces - those all sound very familiar to something else. A computer, a PC! All of these components we find inside the tower of our PC or inside of our laptop. So we have our router, we look inside, we have the same types of components. Let's go through these individually.



For example:

Video Network

The following illustration provides an example of a video network that uses a single Unified Communications Manager cluster. In a successful video network, any endpoint can call any other endpoint. Video availability only exists if both endpoints are video-enabled. Video capabilities extend across trunks.



	The Cisco video conference portfolio comprises the following video bridges: • Cisco TelePresence MCU series • Webex Meeting Server (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Figure 11-14 Cisco WebEx Collaboration Cloud Architecture

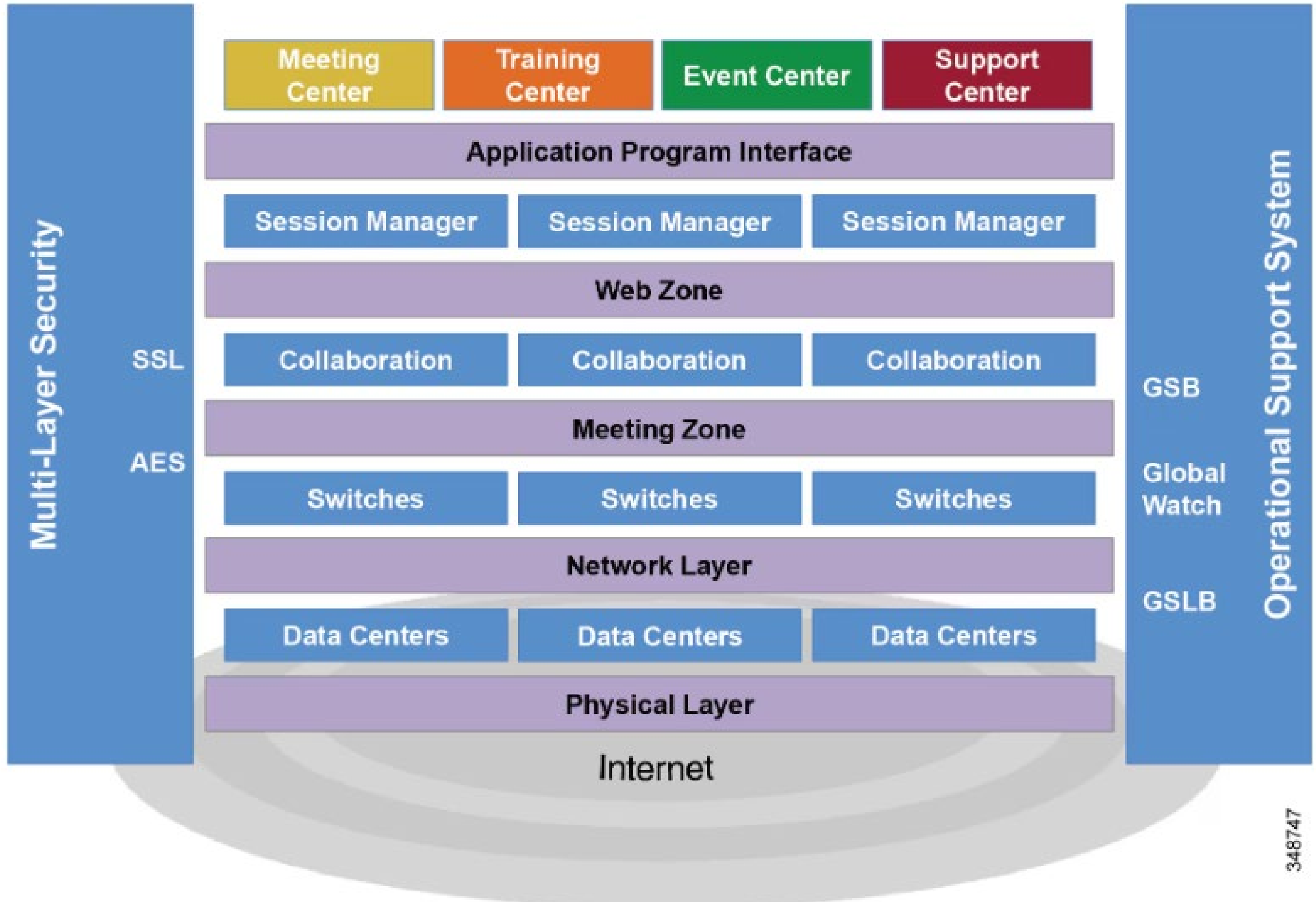
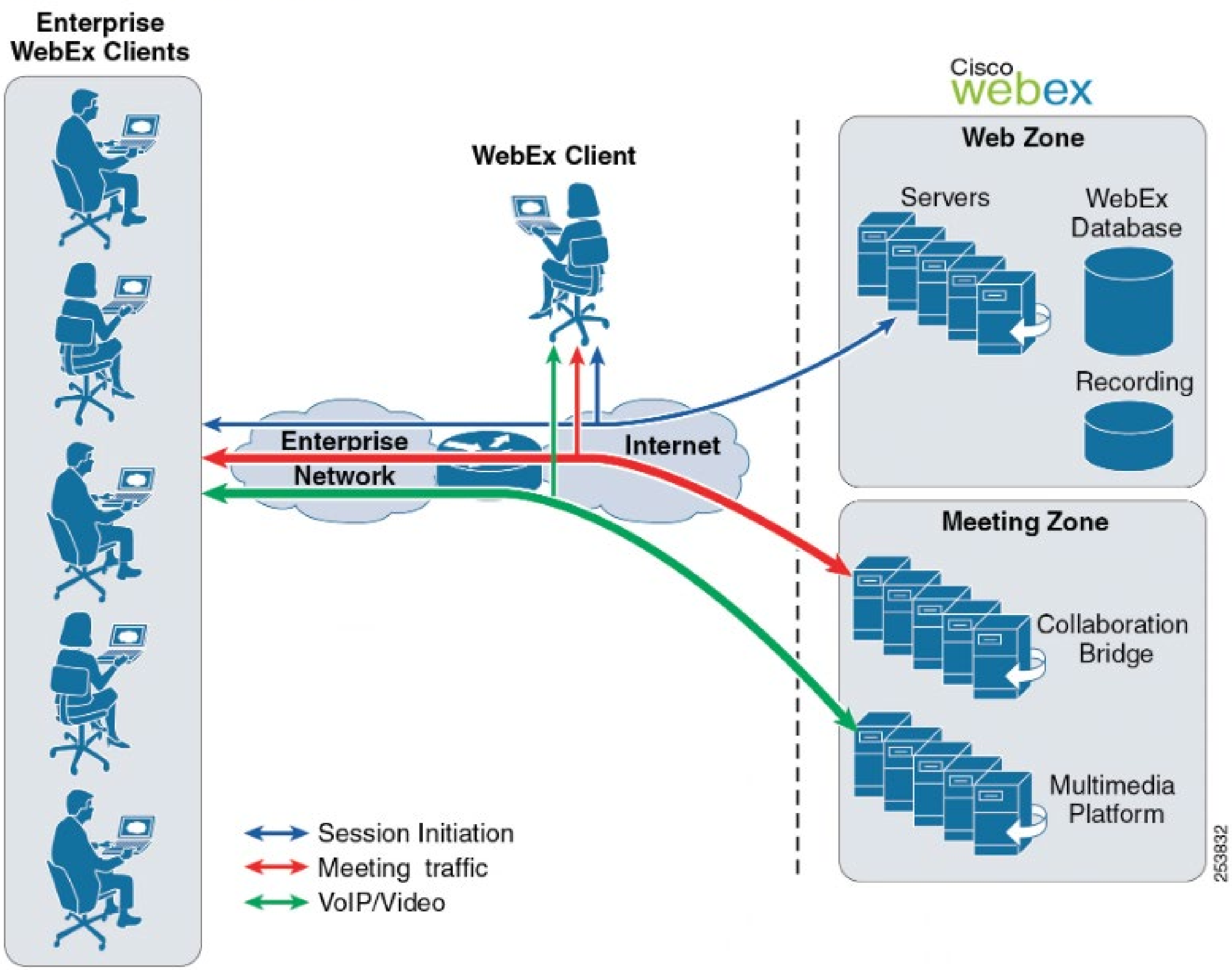


Figure 11-15 WebEx Deployment



(https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)

Product Overview

Cisco Webex Meetings Server is a virtualized, software-based solution that runs on Cisco Unified Computing System™ (Cisco UCS®) servers and VMware. It uses virtual appliance technology for rapid turn-up of services to end users. With Cisco Webex Meetings Server, there are two options for enabling mobile users to more securely access Cisco Webex conferences without going through a VPN. The first option is to deploy reverse proxy (or edge servers) in the enterprise perimeter (or DMZ). The second option, shown in Figure 1, is to deploy the reverse proxy servers behind your internal firewall, thus eliminating all DMZ components and related information security concerns.

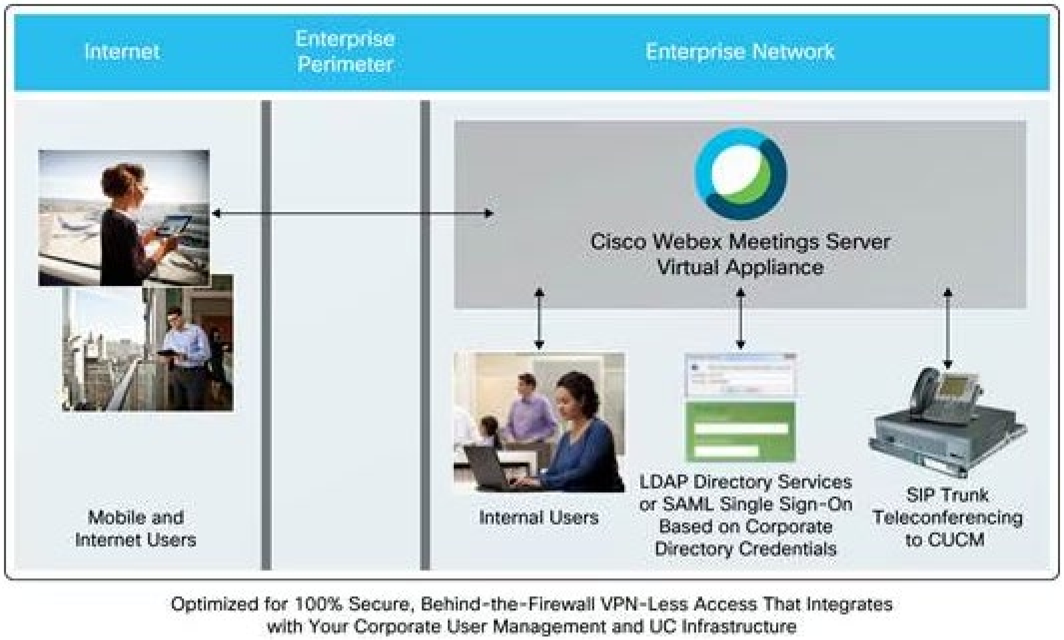


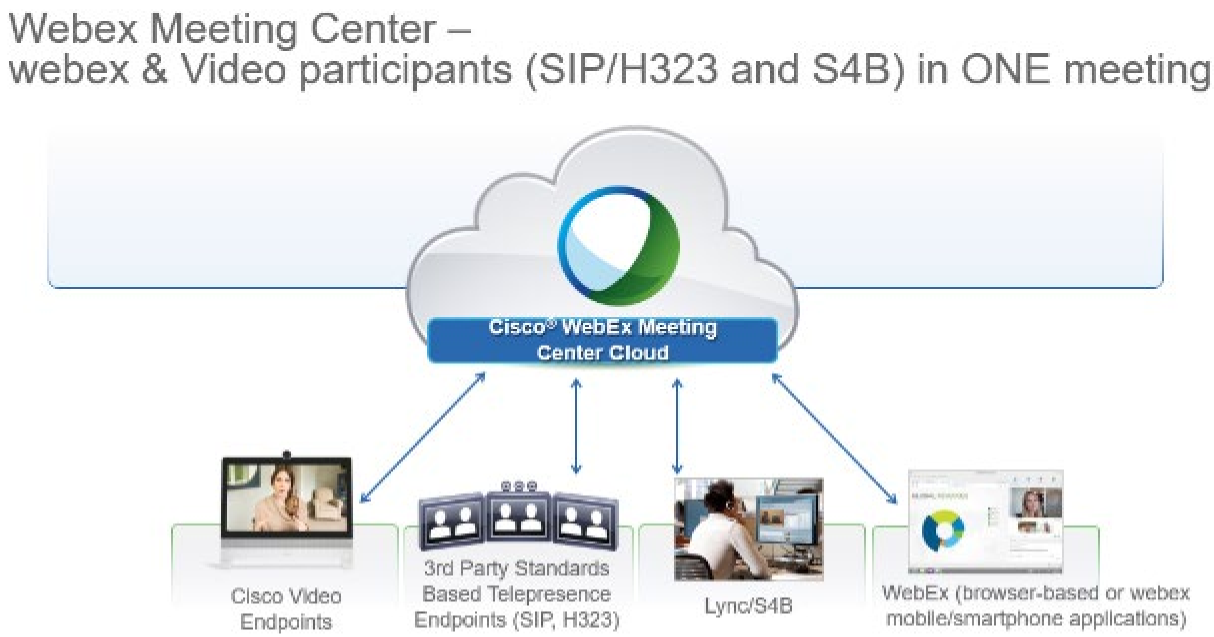
Figure 1.

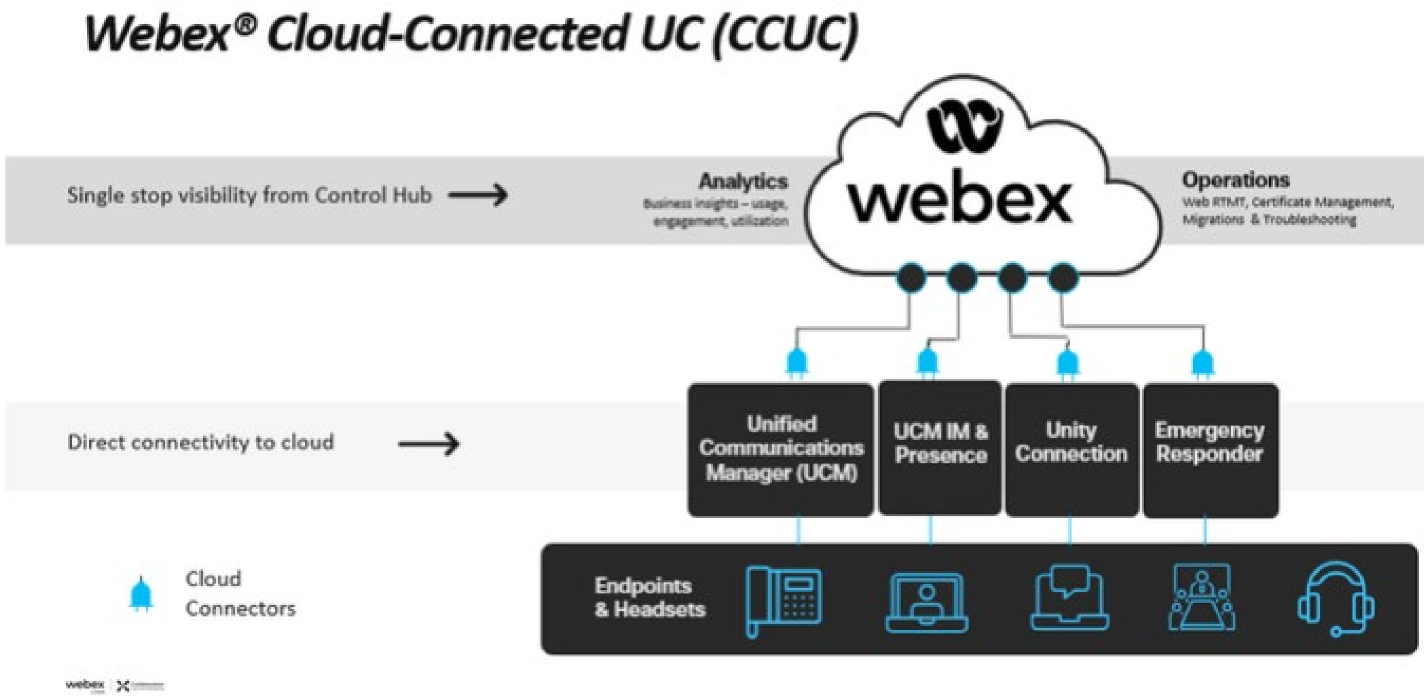
Full Deployment of Cisco Webex Meetings Server Behind a Firewall

System Requirements

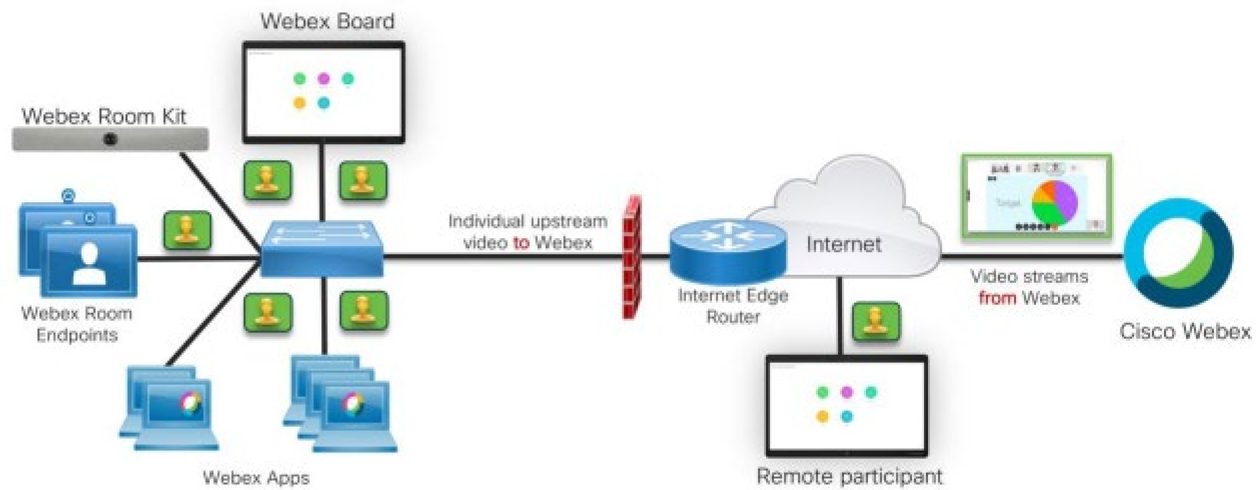
Cisco Webex Meetings Server is compatible with Cisco UCS servers that meet or exceed the specifications presented in this section.

Module	Requirements
Host server	Cisco UCS C-Series rack server or equivalent B-Series blade server
Network interfaces	• Minimum 1 physical Network Interface Card (NIC) for a nonredundant configuration • Redundant configurations must have all NIC interfaces duplicated and connected to an independent switching fabric
Internal storage (direct attached storage [DAS]) for ESXi hosts where internal machines are deployed	Minimum of 4 drives in a RAID-10 or RAID-5 configuration
Internal storage (DAS)for ESXi hosts where Internet Reverse Proxy (IRP) virtual machines are deployed	Minimum of 2 drives in a RAID-1 configuration
SAN storage	Can be used as a substitute for DAS
Network-Attached Storage (NAS)	Can be used as a substitute for DAS or SAN
Hypervisor	• ESXi versions and vSphere licenses • 1 VMware license per processor socket
Email server	• Fully Qualified Domain Name (FQDN) of the mail server that the system uses to send emails • Port number: Default value of the Simple Mail Transfer Protocol (SMTP) port number is 25 or 465 (secure SMTP port number)

	(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings-server/datasheet-c78-717754.html.)  (https://community.cisco.com/t5/collaboration-knowledge-base/cisco-webex-meetings-overview/ta-p/3648888.)

	 (https://community.cisco.com/t5/collaboration-knowledge-base/webex-cloud-connected-uc-connect-your-cucm-to-webex-cloud-for/ta-p/4172103.)
receive a layered video data stream including a base layer and a set of	The Accused Instrumentalities receive a layered video data stream including a base layer and a set of enhancement layers. For example, the Accused Instrumentalities conform to the H.264 standard, including Annex G (SVC). Scalable video routers conforming to the H.264 standard perform a method of receiving a multi-layered stream.

enhancement layers,	sciencedirect.com/topics/engineering/scalable-video-coding **Communicating Pictures: Delivery Across Networks** David R. Bull, in Communicating Pictures, 2014 **Scalable video encoding** It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	 ***Figure 11****: Video Streams in a Webex Meeting* (https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)

https://www.mistralsolutions.com/articles/h-264-svc-update-h-264-video-compression-standard/

+91 80 45621100 info@mistralsolutions.com

About us ▾ Products ▾ Services ▾

MISTRAL®
... Partners in Real Time

AEROSPACE AND DEFENSE
Land, Naval and Airborne

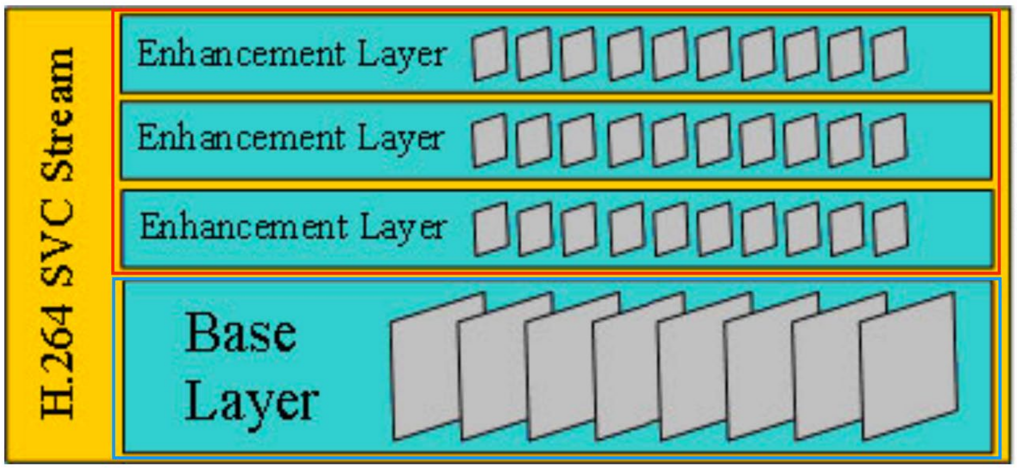


Figure 1: H.264 SVC Layering Technique

The "base layer" is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

G.3.4 **base layer**: A *bitstream subset* that contains all *NAL units* with the nal_unit_type *syntax element* equal to 1 and 5 of the *bitstream* and does not contain any *NAL unit* with the nal_unit_type *syntax element* equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.

See ITU-T H.264 at 437.

https://www.sciencedirect.com/topics/engineering/enhancement-layer 240%

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in The Essential Guide to Video Processing, 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the bitstream into base and enhancement layers. The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, spatial, quality, complexity, and object-based

	Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form "bitstream subsets" that may be decoded as video data streams. **G.3.54 scalable bitstream**: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification. See Telecommunication Standardization Sector, International Telecommunication Union, H.264 Infrastructure of audiovisual services – Coding of moving video, 437 (Jun. 2019) ("ITU-T H.264"). **G.3.6 bitstream subset**: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*. See ITU-T H.264 at 437. Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identify bandwidth-limited conditions of an internet	The Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers. For example, scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between

protocol network between the video router and a plurality of video receivers,	two or more links in the network, a limitation in the capabilities of a given receiver, or some other network condition that would limit the bandwidth between one or more nodes. sciencedirect.com/topics/engineering/scalable-video-coding Communicating Pictures: Delivery Across Networks David R. Bull, in Communicating Pictures, 2014 Scalable video encoding It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	sciencedirect.com/topics/engineering/scalable-coding to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets. For example, the Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers as shown below:

Overview

In this document we will discuss bandwidth utilization. Bandwidth values used will be in payload bit rate which does not include packetization overhead and are covered in 3 categories, average, peak and maximum bit rate:

Average (**avg**) is the average over time for a meeting participant.

Peak (**peak**) is the typical peak bursts over the same time period for a meeting participant.

Maximum (**max**) is the maximum bit rate that the device is capable of either due to device limitations or device configuration.

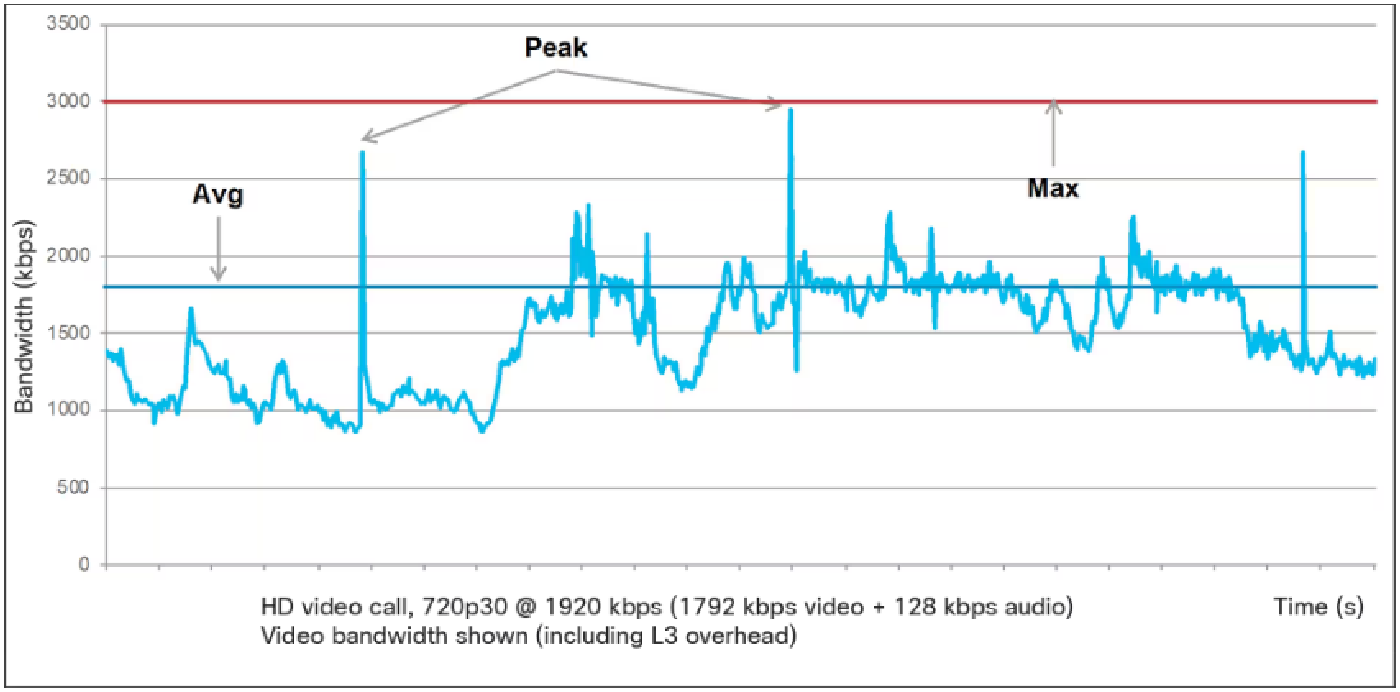
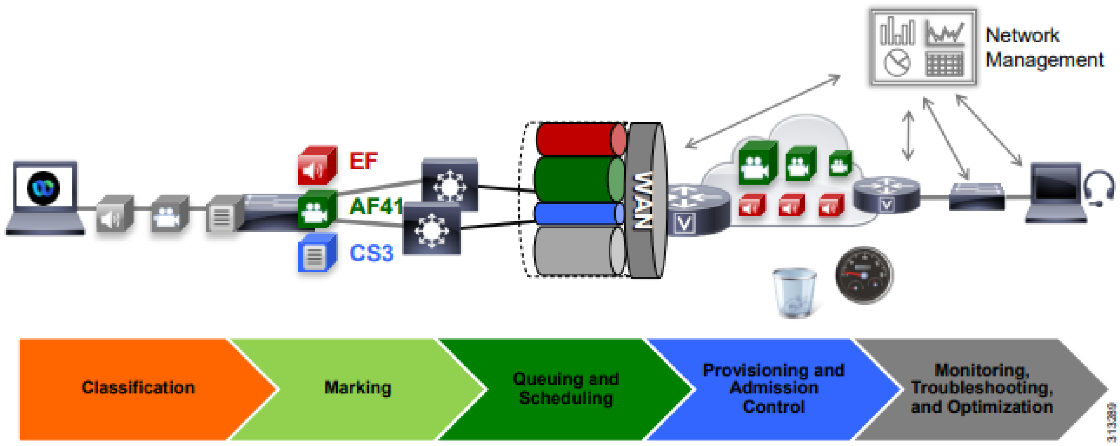


Figure 1.

Video Traffic: Bandwidth Usage High-definition Video Call

(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

Figure 1 Architecture for Bandwidth Management



Recommended Deployment

Modify the existing on-premises QoS switch and WAN and Internet policies to include Webex Services identification, classification, and marking.

- Identify and classify media and signaling traffic from the Webex App and associated workloads as well as Webex Devices.
- Media and signaling marking recommendations:
  1. Mark all audio with Expedited Forwarding class EF (includes all audio of voice-only calls as well as audio for all types of video calls).
  2. Mark all Webex App video with an Assured Forwarding class of AF41 for a prioritized video class of service and AF42 for an opportunistic class of service. The marking of AF41 or AF42 will depend on the choice of whether to deploy opportunistic video during the on-premises deployment phase.
  3. Mark all call signaling with CS3. (All call signaling in HTTPS traffic will be marked based on the enterprise's current policy of traffic marking for HTTP/HTTPS – unless using NBAR which is covered in detail below)
- Configure QoS on all media originating and terminating applications such as the Video Mesh Nodes, Expressway and Cisco Unified Border Element.
- Update the WAN edge ingress re-marking policy.
- Update the WAN edge egress queuing and scheduling policy.

Every Cisco video endpoint employs several smart media techniques to avoid network congestion, recover from packet loss, and optimize network resources. The following smart media techniques are just some of the techniques employed by Cisco Webex App and Devices:

Media resilience techniques

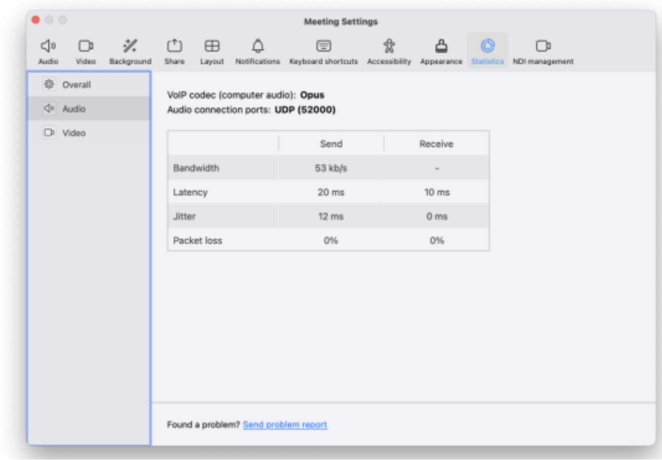
- Encoder pacing
- Forward Error Correction (FEC)
- Rate adaptation

(https://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Collaboration/AltDesigns/BWM-Wbx.pdf.)

Poor Audio / Video Quality - Full-featured Meetings

Help > Health Checker > Audio and Video Statistics...

- Indicates TCP or UDP w/ Source Port
- Latency / Packet Loss / Jitter



(https://www.ciscolive.com/c/dam/r/ciscolive/global-event/docs/2024/pdf/BRKCOL-3431.pdf.)

There are three main factors that impact the quality of an audio or video call. These factors are **packet loss**, **latency**, and **jitter**. As shown in Figure 2, **packet loss is simply losing one or more packets within a stream of packets**. In this example, a Voice over IP (VoIP) packet is lost going from the Webex client to the Webex server. Loss can occur for a number of reasons in a network but most commonly it is caused by congestion or resource contention in the network or on the endpoints. This leads to routers, switches, or the endpoints themselves dropping or delaying packets. In other cases, packets encounter such a high delay that they are received too late to be played out. These packets will also be counted as lost.

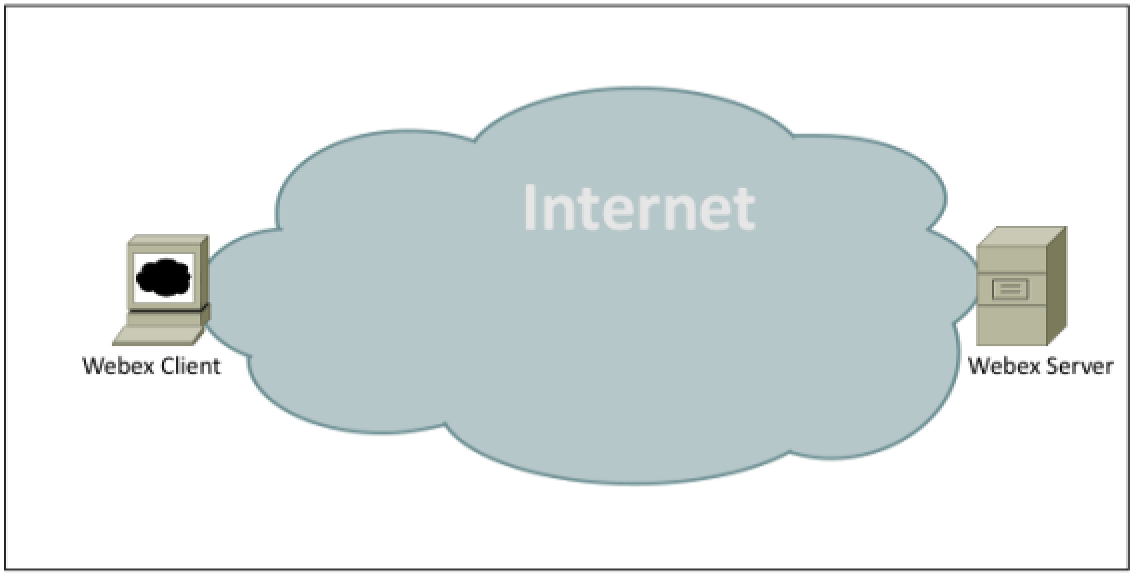
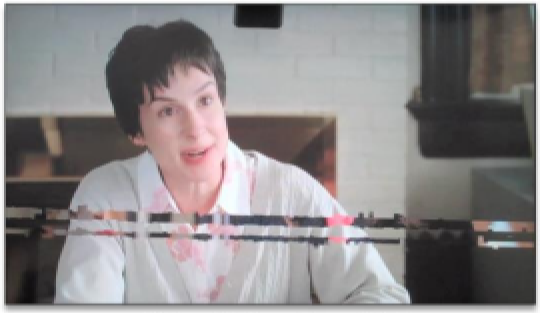


Figure 2: *Packet Loss Example*

Note: *Control Hub Diagnostics highlights potential problems with packet loss and delay by coloring voice and video streams as Good, Fair, or Poor. It is important to understand that just because Control Hub Diagnostics flags part of a conversation as Fair or Poor, it does not necessarily mean that the user had a bad experience. Voice and video compensation mechanisms and algorithms in combination with other factors can mitigate the impact of loss and delay to the point where it is not noticeable by the user.*

Video Quality Artifacts

The first step in troubleshooting video quality is to clearly understand the symptom experienced by the end user. Getting a screenshot or a picture of the display screen usually gives more information than a verbal description of the problem. The type of video artifacts observed in a screenshot or picture can be used to determine the troubleshooting path. For example, in Figure 13, pictures (a) and (b) are artifacts that are caused when video streams experience packet loss. Just by seeing the picture, you can immediately start troubleshooting to identify the source of the packet loss.



a. Video with line stripe artifacts

b. Frozen video with block artifacts

Webex Control Hub makes it quite easy to find the device type and operating system version used by the end user to join a Webex meeting. All you need is the end user's email address and meeting time to find the right meeting and to get end user's device details. Figure 14 shows the list of meetings attended by a participant with the email address, rtpmsuser1@gmail.com.

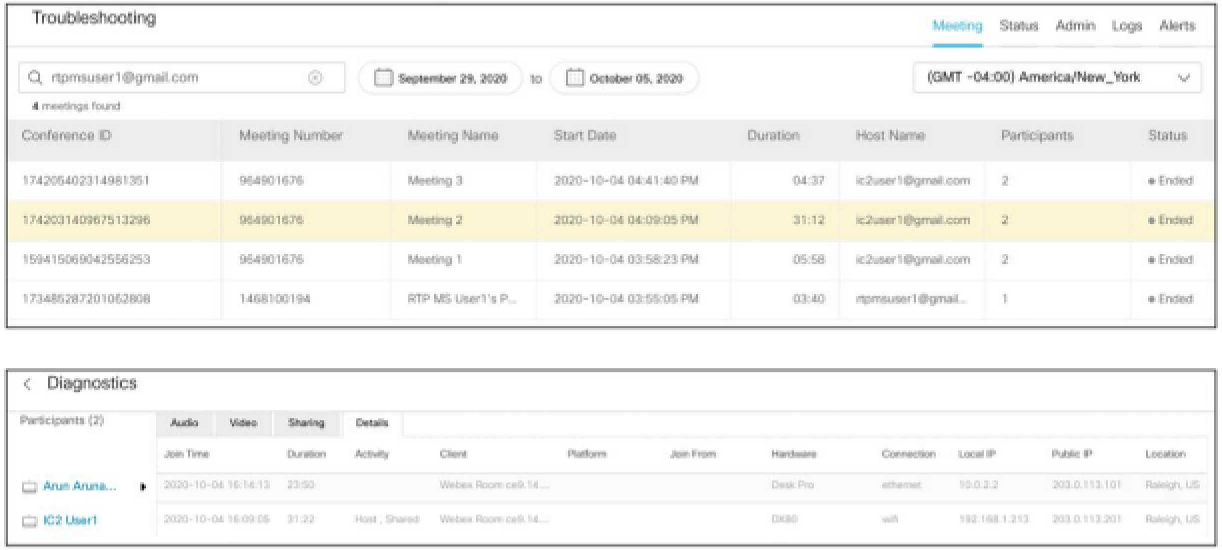


Figure 14*: Participant Device Details*

(https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)

forward the base layer from the video router to at	The Accused Instrumentalities forward the base layer from the video router to at least two of the plurality of video receivers via the internet protocol network, and selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions.

least two of the plurality of video receivers via the internet protocol network, and selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet	For example, because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer. **G.8.8.2 Specification of the base layer bitstream** Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. **Rec. ITU-T H.264 (06/2019)** 595 If a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth-limiting condition, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For instance, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.

protocol network based upon the identified bandwidth-limited conditions,	The sub-bitstream is derived by applying the following operations in sequential order: 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 489.

The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation.

See ITU-T H.264 at 624.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

	For example, the Accused Instrumentalities selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions:

Table 4. Webex Meetings Bandwidth per Resolution Table

Layer	Bandwidth Range
90p active thumbnail (each)	~60-100 kb/s
180p main video	125-200 kb/s
360p main video	470-640 kb/s
720p main video	900k-1.5 mb/s
Content sharing (sharpness, 1080p/5)	120k – 1.3 mb/s
Content sharing (motion, 720p/30)	900k – 2.5 mb/s

Webex Meetings Desktop App Bandwidth Controls

Webex administrators have 2 key controls to help control bandwidth as used by clients that connect to Webex meetings should they choose to. Namely, you can cap the meeting layouts at either 360p as the max available resolution, or to enable 720p layers. Whether your site is administered on Webex Control Hub or Webex Site Administrator, the following controls are available in Configuration > Common Site Settings > Options:

☐ Turn on high-quality video (360p) *(Meetings, Training, Events and Support)*

☐ Turn on high-definition video (720p) *(Meetings, Training and Events)*

Figure 5.

Webex Meetings Desktop App Bandwidth Controls

Webex Media Improvements

The following are media improvements that have occurred in releases from 40.7 – 40.10

Improved Video Experience over WIFI (Defer video down-speeding): WIFI environments present a specific set of challenges and through testing we noticed the algorithms were too quick to drop resolution/bandwidth in certain WIFI environments. In 40.7 the algorithms have been updated to 'defer the down-speeding" of video for brief WIFI specific network dropouts. This resulted in improved user experience over WIFI. See figure5 for an illustration of this. When latency spikes for a specific duration the improved algorithms can detect that this is a WIFI specific event and not thus delay the down-speeding of the video resolution until that duration (n) is exceeded in which case would infer a different network event increasing latency.

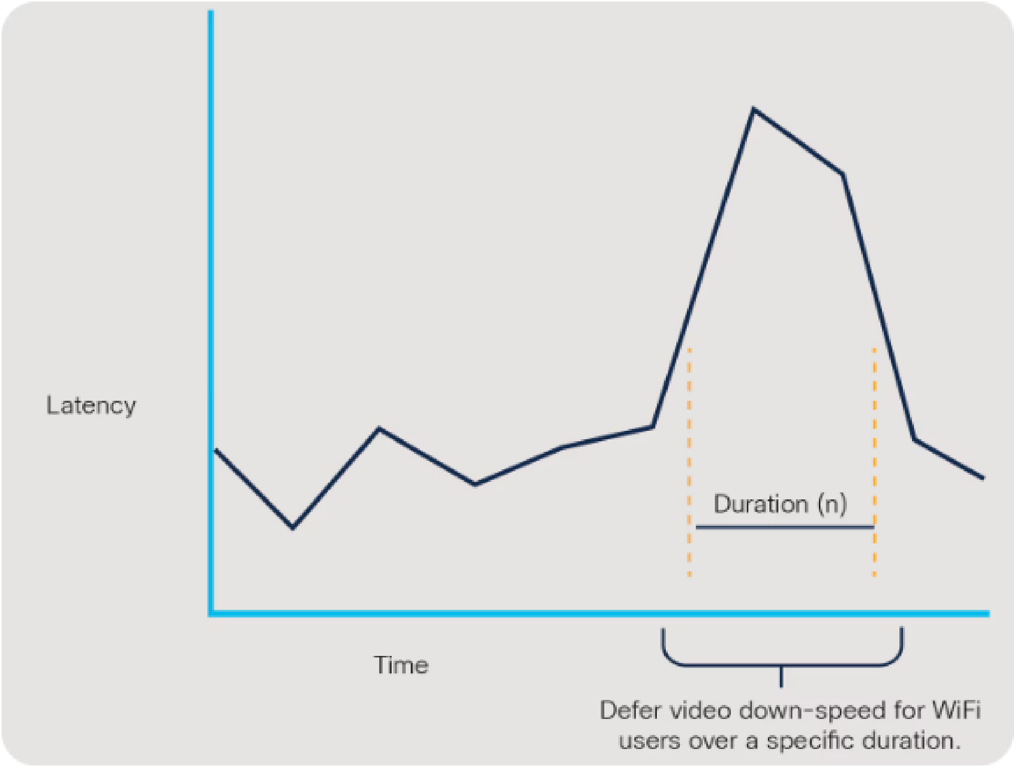


Figure 6.

Deferred Video Down-speeding

Video Super Scaling is a technique used for users on sub-optimal networks. We provide 720p like experiences to users who are experiencing network impairments resulting in 360p connections to our servers for their meeting. Our updated scaling algorithm provides enhanced video quality when this degradation occurs, making network impairment downgrades less quality impacting. If a 360p image is simply scaled geometrically to a 720p resolution, the quality has not improved, but our technique involves migrating video post-processing tasks to the client for some video flows versus doing it in the cloud, resulting in being able to keep more precious bits of information to reconstruct the video at a given bandwidth.

In a meeting as shown in figure 8 and 9, the bandwidth requirements are as follows. Assuming the receiver is not bandwidth restricted, Webex app will adapt to available network conditions. Table 8 outlines the bandwidth utilization. Figure 13 illustrates the Active Speaker layout which is the default meeting layout for Webex Teams App.

Table 8. Cisco Webex Teams Bandwidth Utilization

Meeting Scenario	Main Video+Audio (no content sharing)	Main Video+Audio (content sharing)
Webex Teams 1:1 Call	1.5 mb/s peak, 1.3 mb/s avg	1.5 mb/s peak, 1.3 mb/s avg
Webex Teams Meeting (fig 13 layout), 6 participants	2.3 mb/s peak, 1.7 mb/s avg	2.9/ mb/s peak, 2.0 mb/s avg
Webex Teams Meeting (fig 14 layout), 6 participants	2.5 mb/s peak, 1.8 mb/s avg	3.5 mb/s peak, 2.3 mb/s avg

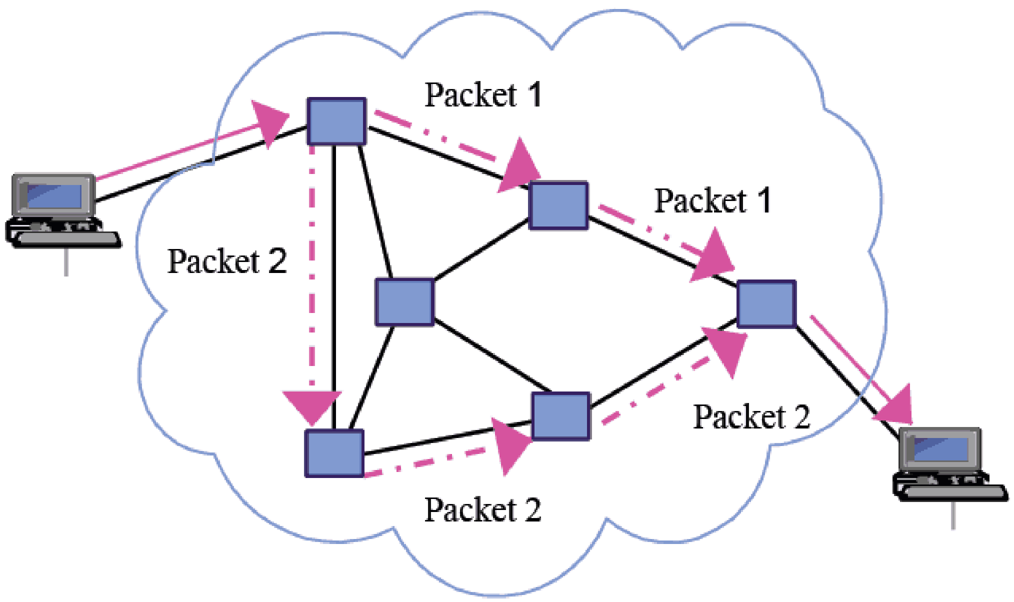
(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

and wherein the video router transmits the layered video data stream according to an internet protocol; wherein each layer of the layered video data stream comprises data packets,	The Accused Instrumentalities include a video router, wherein the video router transmits the layered video data stream according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets. For example, scalable video routers facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.

networkencyclopedia.com/packet-switching/

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or "packets" that contain sequence numbers so that they can be reassembled at the destination.



Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

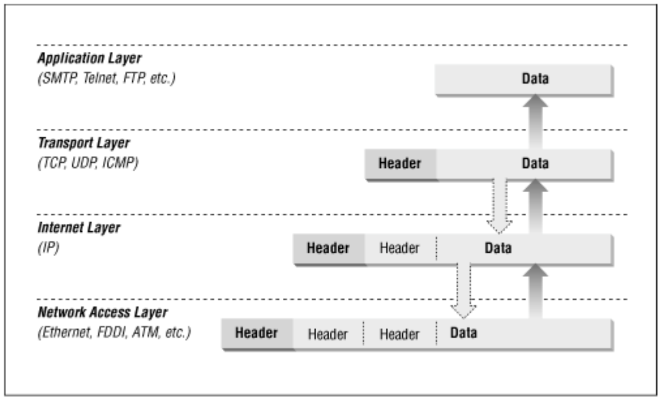
Not Secure | web.deu.edu.tr/doc/oreily/networking/firewall/ch06_03.htm

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. Figure 6.2 shows how this works.

Figure 6.2: Data encapsulation

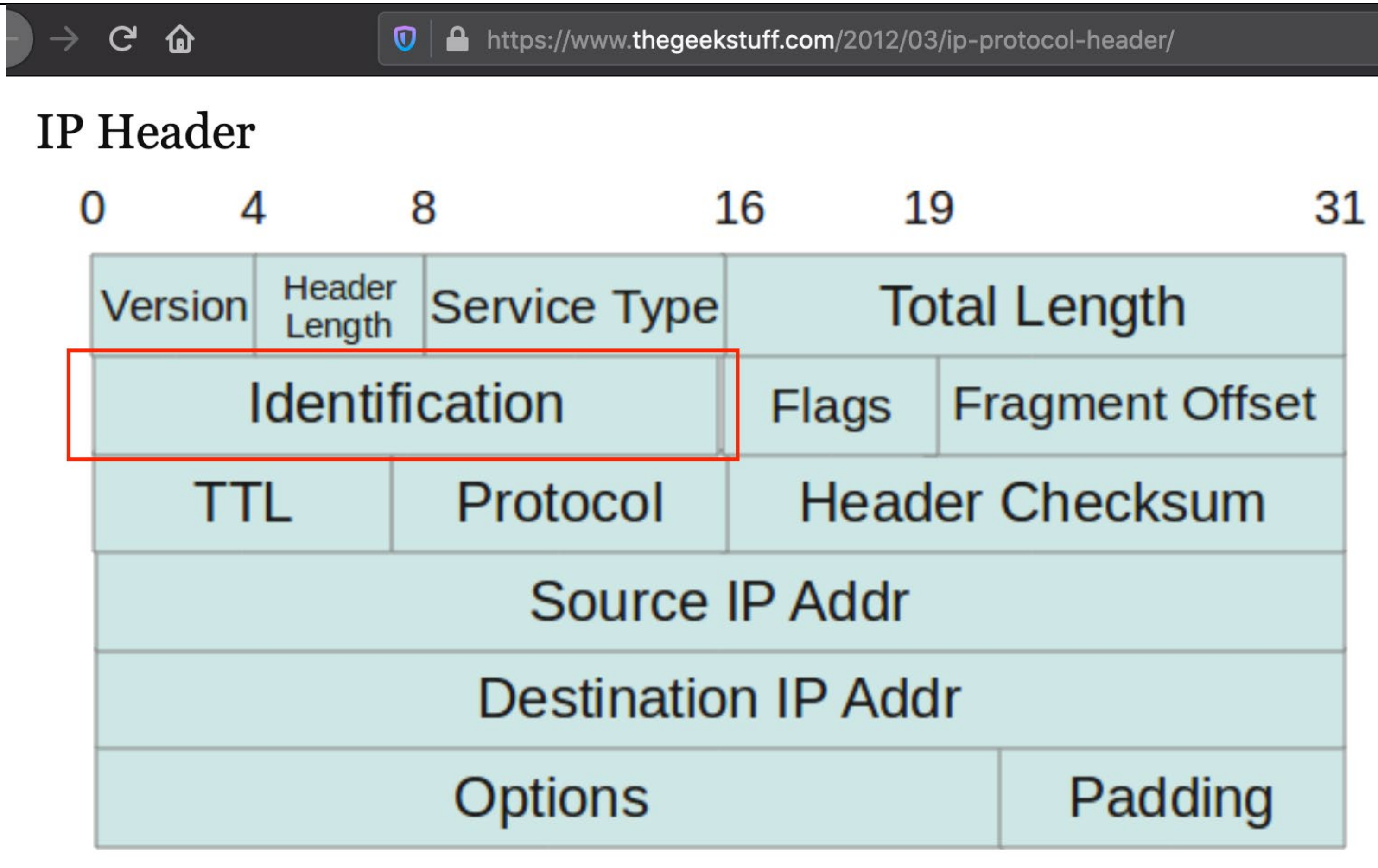


At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

	ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf *A. Network Abstraction Layer (NAL)* The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-
each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a	The Accused Instrumentalities include a video router, wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. For example, each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence.

layer to which the packet belongs.	Not Secure | linfo.org/packet_header.html **Packet Header Definition** A *packet header* is the portion of an IP (Internet protocol) *packet* that precedes its body and contains addressing and other data that is required for it to reach its intended destination. Packets are the fundamental unit of information transport in all modern computer networks, and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the *payload*, and a *trailer*. The header's format is specified in the Internet protocol. It normally contains 20 bytes of data, although an option exists within it that allows the addition of more bytes. Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's IP address, the intended receiver's IP address, the number of packets the message has been broken into, the identification number of the particular packet, the protocol (e.g. 1 for ICMP, 2 for IGMP, 6 for TCP and 17 for UDP) used, the packet length (on networks that have variable length packets), the *time to live* (i.e., the number of links or *hops* that the packet can be routed before being allowed to expire) and synchronization data (several bits that help the packet match up to the network).


https://www.thegeekstuff.com/2012/03/ip-protocol-header/
IP Header
0
4
8
16
19
31
Version
Header Length
Service Type
Total Length
Identification
Flags
Fragment Offset
TTL
Protocol
Header Checksum
Source IP Addr
Destination IP Addr
Options
Padding

https://www.thegeekstuff.com/2012/03/ip-protocol-header/

- ***Identification(16 bits)***: This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

For example, each data packet is encoded with a layer identifier, wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.

Each scalable layer is associated with a unique layer identifier as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier layerId does not include any scalable layer with a layer identifier greater than layerId, but it may include scalable layers with layer identifiers less than layerId. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause.

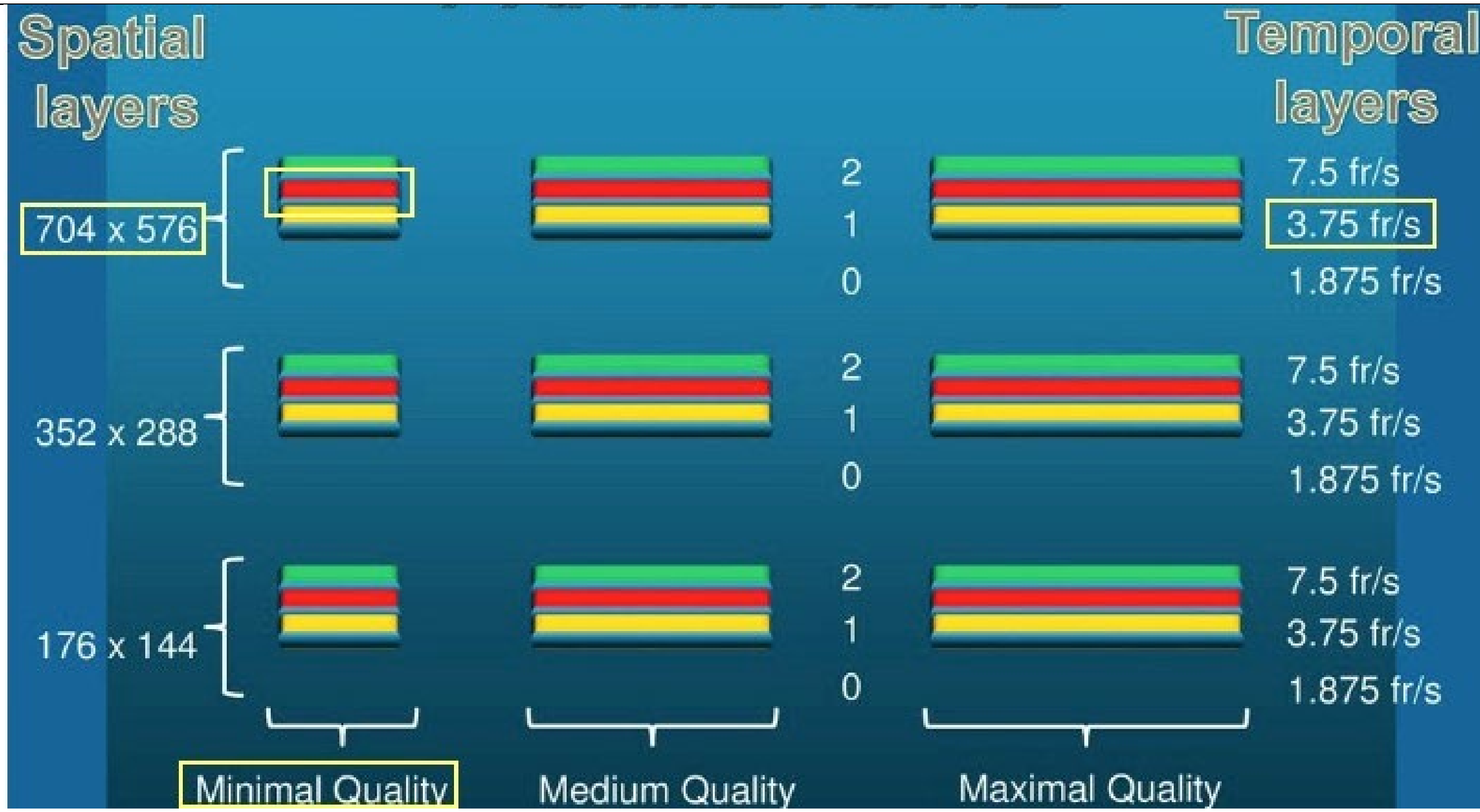
NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have sub_pic_layer_flag[i] equal to 0, the unique layer identifier values may be set equal to (128 * dependency_id + 8 * quality_id + temporal_id), with dependency_id, quality_id, and temporal_id being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer.

See ITU-T H.264 at 624.

NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including priority_id dependency_id, quality_id, and temporal_id, according to the standard.

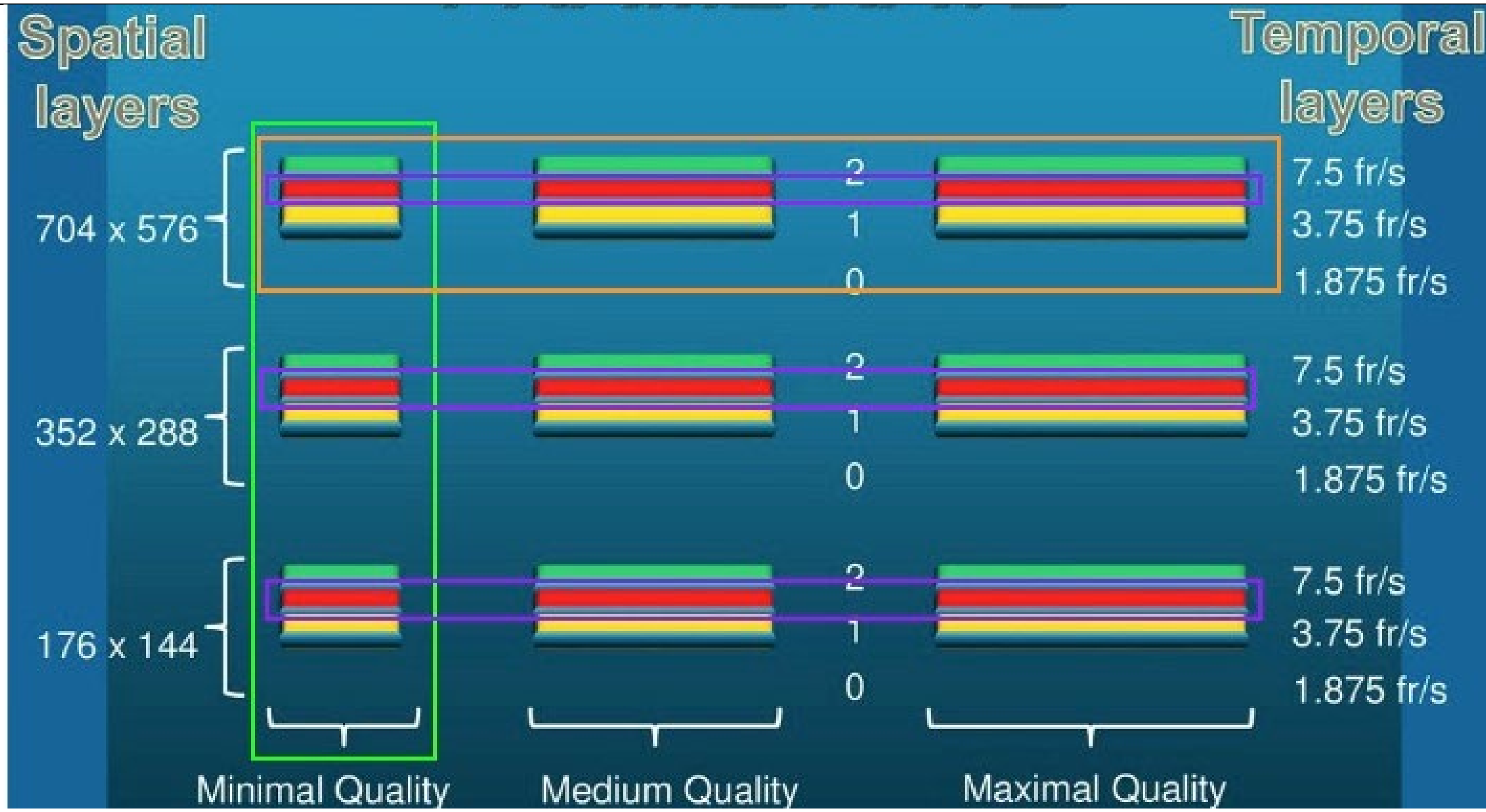
dependency_id specifies a dependency identifier for the NAL unit. dependency_id shall be equal to 0 in prefix NAL units. The assignment of values to dependency_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1.

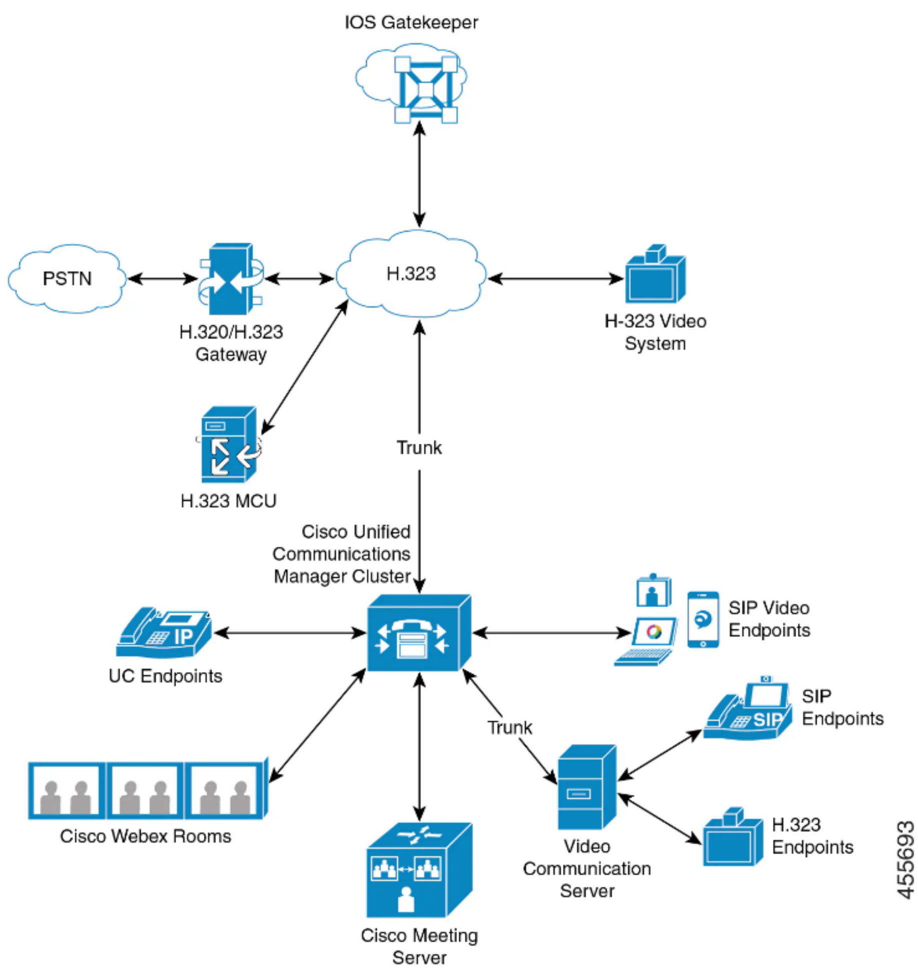
	See ITU-T H.264 at 459. **priority_id** specifies a priority identifier for the NAL unit. The assignment of values to priority_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458. **quality_id** specifies a quality identifier for the NAL unit. quality_id shall be equal to 0 in prefix NAL units. The assignment of values to quality_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable DQId is derived by $DQId = (dependency\_id << 4) + quality\_id$ (G-63) When nal_unit_type is equal to 20, the bitstream shall not contain data that result in DQId equal to 0. See ITU-T H.264 at 459. **temporal_id** specifies a temporal identifier for the NAL unit. The assignment of values to temporal_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The value of temporal_id shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with nal_unit_type equal to 5 or idr_flag equal to 1, temporal_id shall be equal to 0. See ITU-T H.264 at 459. For instance, the graphic below illustrates an example of a hierarchal virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.



Source: https://slideplayer.com/slide/13389021/.

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.

	 Source, https://slideplayer.com/slide/13389021/.
6. A method for transmitting video signals, the method comprising:	The Accused Instrumentalities perform a method for transmitting video signals. Cisco performs the claimed method, for example, by using the Accused Instrumentalities in its provision of monthly subscription services to Webex. For example, scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding IP data packets containing scalable video information.

	Video Network The following illustration provides an example of a video network that uses a single Unified Communications Manager cluster. In a successful video network, any endpoint can call any other endpoint. Video availability only exists if both endpoints are video-enabled. Video capabilities extend across trunks. 

	The Cisco video conference portfolio comprises the following video bridges: • Cisco TelePresence MCU series • Webex Meeting Server (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Figure 11-14 Cisco WebEx Collaboration Cloud Architecture

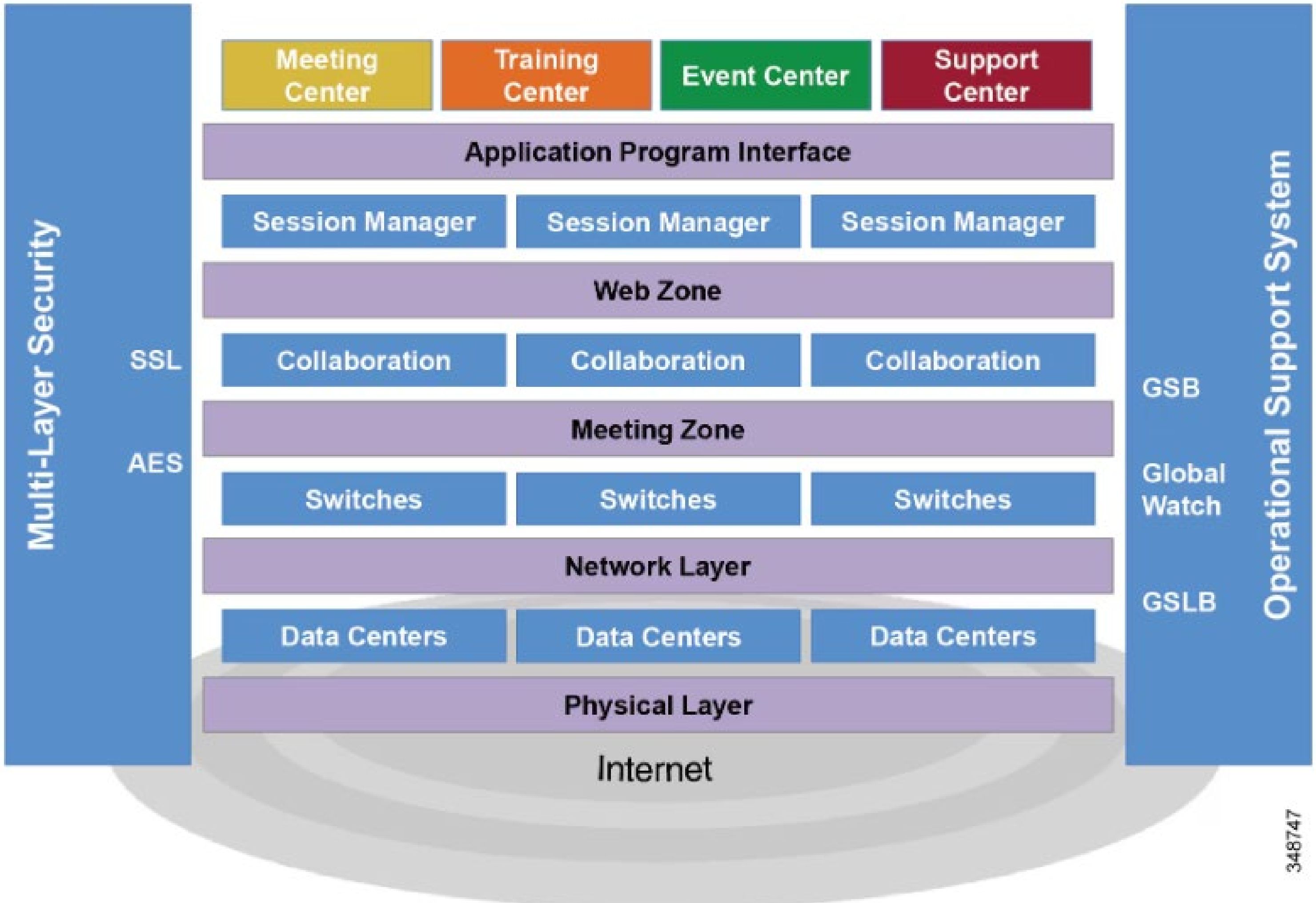
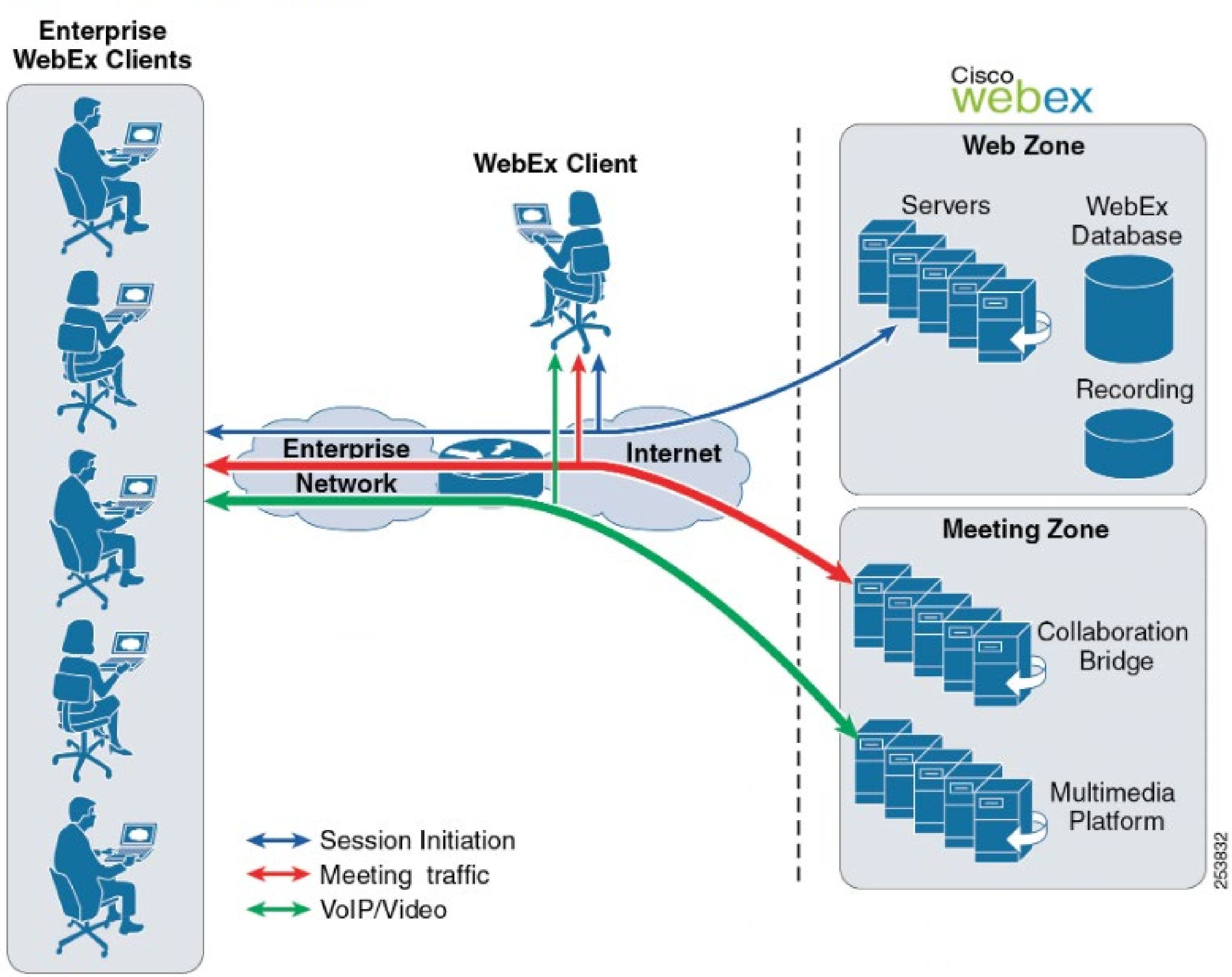


Figure 11-15 WebEx Deployment



	Cisco Rich Media Conferencing consists of the conferencing solutions described below. The details pertaining to each solution are described in each individual section that follows. • Cisco Unified CM Audio Conferencing This solution allows Unified CM to use its internal software component or external hardware digital signal processors (DSPs) as the resources to perform audio conferencing. • Cisco Meeting Server Cisco Meeting Server is an on-premises video conferencing solution. Each user has a personal Space that can be used to conduct meetings. Users can manage items such Space creation, adding members to a Space, and PIN creation from the Cisco Meeting App. • Cisco Collaboration Meeting Rooms Hybrid Cisco CMR Hybrid combines the on-premises video conference and the WebEx Meeting Center conference into a single meeting, which allows TelePresence and WebEx participants to join and share voice, video, and content. CMR Hybrid meetings can be either scheduled or non-scheduled. • Cisco WebEx Meeting Center Video Conferencing Cisco WebEx Meeting Center Video Conferencing (formerly Cisco Collaboration Meeting Rooms (CMR) Cloud) is an alternate conferencing deployment model that does not require any on-premises conferencing resources or management infrastructure. It supports both scheduled and non-scheduled meetings as well as TelePresence, audio, and WebEx participants in a single call, all hosted in the cloud. • Cisco WebEx Meetings Server Where cloud-based web and audio conferencing is not suitable, it is possible to use the on-premises WebEx Meetings Server solution. This product offers a standalone audio, video, and collaboration web conferencing platform. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)

The Accused Instrumentalities implement the scalable video coding features in Annex G of the H.264 standard. For example:

Video Calls

The typical video call includes two or three Real-Time Protocol (RTP) streams in each direction (that is, four or six streams). The call can include the following stream types:

- Video (H.261, H.263, H.263+, H.264-SVC, X-H.264UC, H.264-AVC, H.265, AV1 and VT Camera wideband video codecs)
- Far-End Camera Control (FECC) - Optional
- Binary Floor Control Protocols (BFCP)

SIP Video

SIP video supports the following video calls by using the SIP Signaling Interface (SSI):

- SIP to SIP
- SIP to H.323
- SIP to SCCP
- SIP intercluster trunk
- H.323 trunk
- Combination of SIP and H.323 trunk

SIP video calls also provide media control functions for video conferencing.

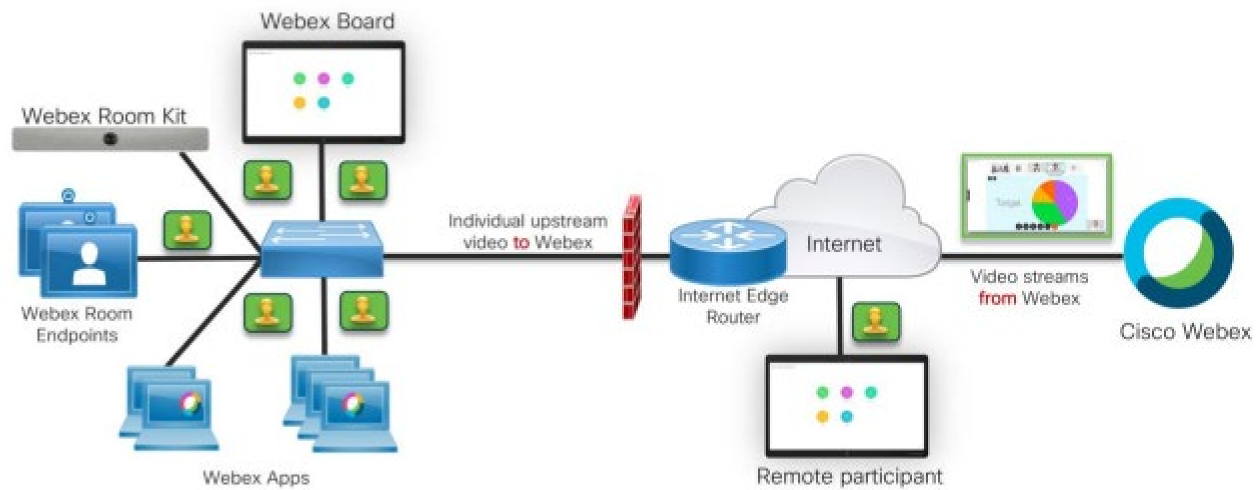
Unified Communications Manager video supports SIP on both SIP trunks and lines support video signaling. SIP supports the H.261, H.263, H.263+, H.264 (AVC), H.264 (SVC), X-H.264UC (Lync), and AV1 video codecs (it does not support the wideband video codec that the VTA uses).

	Video Codecs Common video codecs include H.261, an older video codec, H.263, a newer codec that gets used to provide internet protocol (IP) video, and H.264, a high-quality codec. The system supports H.264 for calls that use the Skinny Client Control Protocol (SCCP), H.323, and SIP on originating and terminating endpoints only. The system also supports regions and locations. Unified Communications Manager maintains the offerer's video codec ordering preference when making the answer, if possible. H.265 is the preferred video codec is available on the endpoints, otherwise, Unified Communications Manager follows the following codec preference preference order: (table below) (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Preference Order	Codecs	Description
1	H.265 (HEVC)	Provides higher quality video using lower bandwidth.
2	H.264 (SVC)	Allows rendering of variable quality video from the same media stream, by disregarding a subset of the packets received. **Note** H. 264 SVC is a new annex to the H.264-AVC video compression standard; meaning it is an enhancement on top of H.264-AVC. It provides the ability to encapsulate multiple video streams at various frame-rates and resolutions in one container.

	Meeting Center uses the H.264 AVC/SVC codec to provide high-definition video for the conference. Higher network bandwidth is needed for those deployments. For further details regarding network traffic optimization for high-definition video, see Capacity Planning. **Network Traffic Planning** Network traffic planning for Cisco WebEx Meeting Center Video Conferencing consists of the following elements: • WebEx Clients bandwidth The WebEx meeting client uses Scalable Video Coding (SVC) technology to send and receive video. It uses multi-layer frames to send video, and the receiving client automatically selects the best possible resolution to receive video that typically requires 1.2 to 3 Mbps of available bandwidth. For more information regarding network traffic planning for WebEx clients, refer to the *Cisco WebEx Network Bandwidth* white paper, available at https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meeting-center/white_paper_c11-691351.html • Bandwidth for video device from enterprise to WebEx Cloud For optimal SIP audio and video quality, Cisco recommends setting up the video bandwidth for at least 1.5 Mbps per device screen in the region associated with the endpoint registering with Cisco Unified CM. For example, if a triple-screen device registers with Unified CM, video bandwidth of 4.5 Mbps should be allocated in the associated region. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)
receiving a layered video data stream comprising a base layer	The Accused Instrumentalities perform a method for transmitting video signals, comprising receiving a layered video data stream comprising a base layer and a set of enhancement layers. For example, the Accused Instrumentalities conform to the H.264 standard, including Annex G (SVC). Scalable video routers conforming to the H.264 standard perform a method of receiving a multi-layered stream.

and a set of enhancement layers;	sciencedirect.com/topics/engineering/scalable-video-coding **Communicating Pictures: Delivery Across Networks** David R. Bull, in Communicating Pictures, 2014 **Scalable video encoding** It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	 ***Figure 11**: Video Streams in a Webex Meeting* (https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)

https://www.mistralsolutions.com/articles/h-264-svc-update-h-264-video-compression-standard/

+91 80 45621100 info@mistralsolutions.com

About us ▾ Products ▾ Services ▾

MISTRAL ®
... Partners in Real Time

AEROSPACE AND DEFENSE
Land, Naval and Airborne

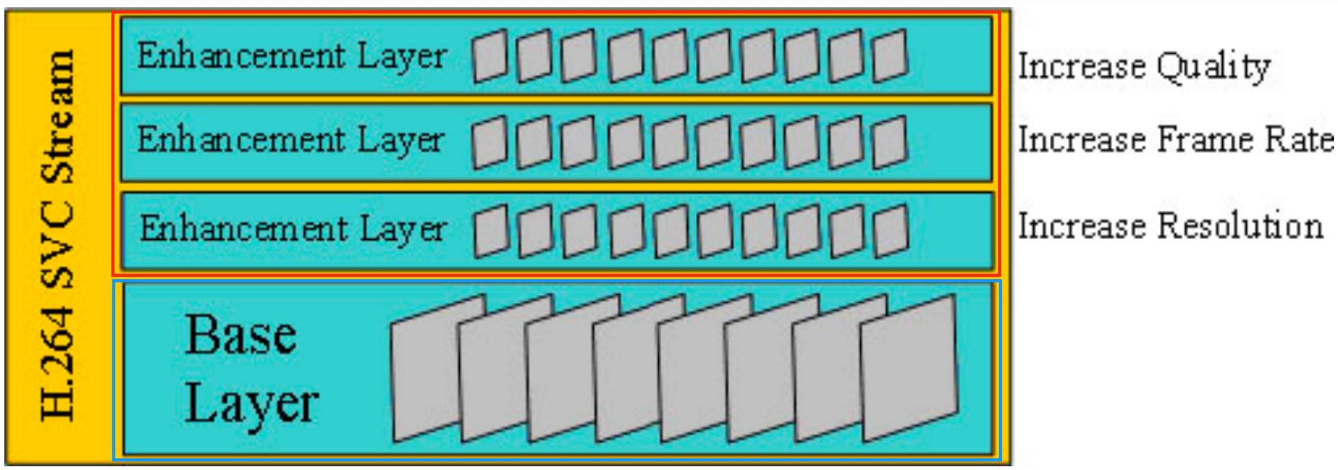


Figure 1: H.264 SVC Layering Technique

The "base layer" is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

> **G.3.4** **base layer**: A *bitstream subset* that contains all *NAL units* with the nal_unit_type *syntax element* equal to 1 and 5 of the *bitstream* and does not contain any *NAL unit* with the nal_unit_type *syntax element* equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.

See ITU-T H.264 at 437.

https://www.sciencedirect.com/topics/engineering/enhancement-layer

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in The Essential Guide to Video Processing, 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the bitstream into base and enhancement layers. The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, spatial, quality, complexity, and object-based

	Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form "bitstream subsets" that may be decoded as video data streams. **G.3.54 scalable bitstream**: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification. See Telecommunication Standardization Sector, International Telecommunication Union, H.264 Infrastructure of audiovisual services – Coding of moving video, 437 (Jun. 2019) ("ITU-T H.264"). **G.3.6 bitstream subset**: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*. See ITU-T H.264 at 437. Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identifying bandwidth-limited conditions of an internet protocol	The Accused Instrumentalities perform a method for transmitting video signals, comprising identifying bandwidth-limited conditions of an internet protocol network between a video router and a plurality of video receivers. For example, scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between

network between a video router and a plurality of video receivers;	two or more links in the network, a limitation in the capabilities of a given receiver, or some other network condition that would limit the bandwidth between one or more nodes. sciencedirect.com/topics/engineering/scalable-video-coding **Communicating Pictures: Delivery Across Networks** David R. Bull, in Communicating Pictures, 2014 **Scalable video encoding** It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

For example, the Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers as shown below:

Overview

In this document we will discuss bandwidth utilization. Bandwidth values used will be in payload bit rate which does not include packetization overhead and are covered in 3 categories, average, peak and maximum bit rate:

Average (**avg**) is the average over time for a meeting participant.

Peak (**peak**) is the typical peak bursts over the same time period for a meeting participant.

Maximum (**max**) is the maximum bit rate that the device is capable of either due to device limitations or device configuration.

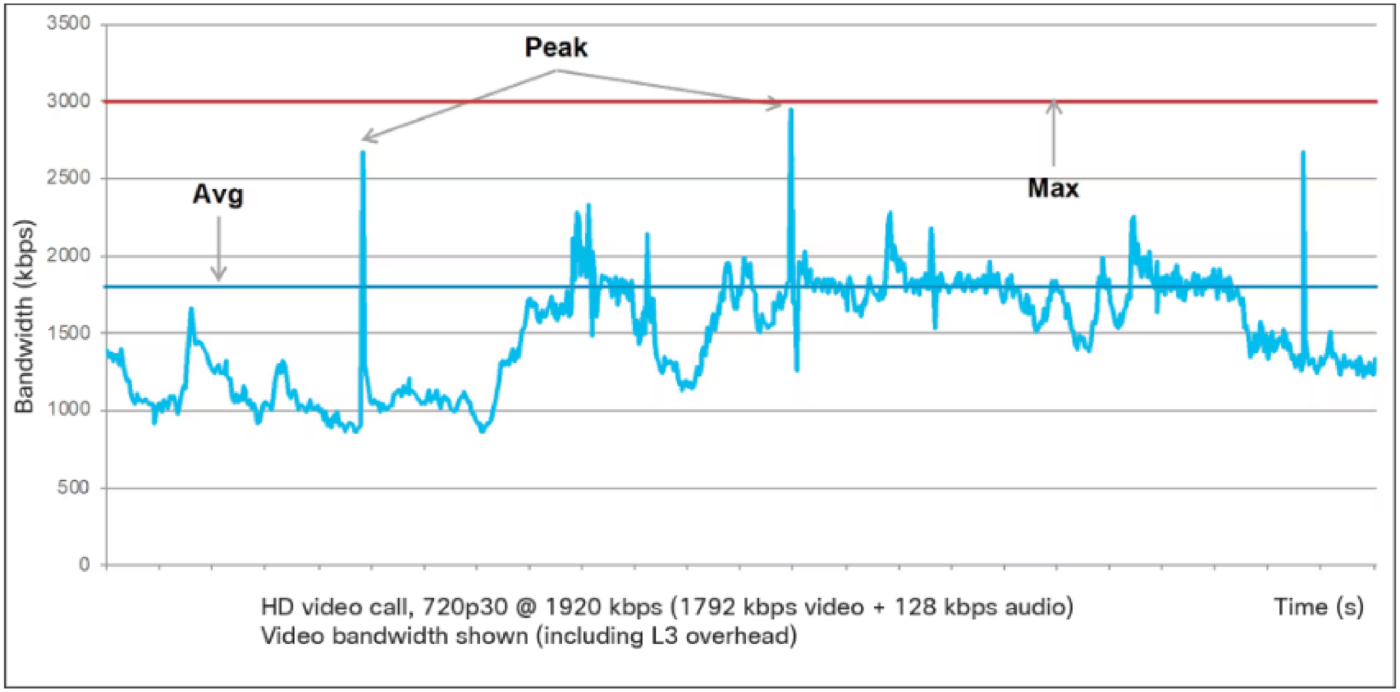
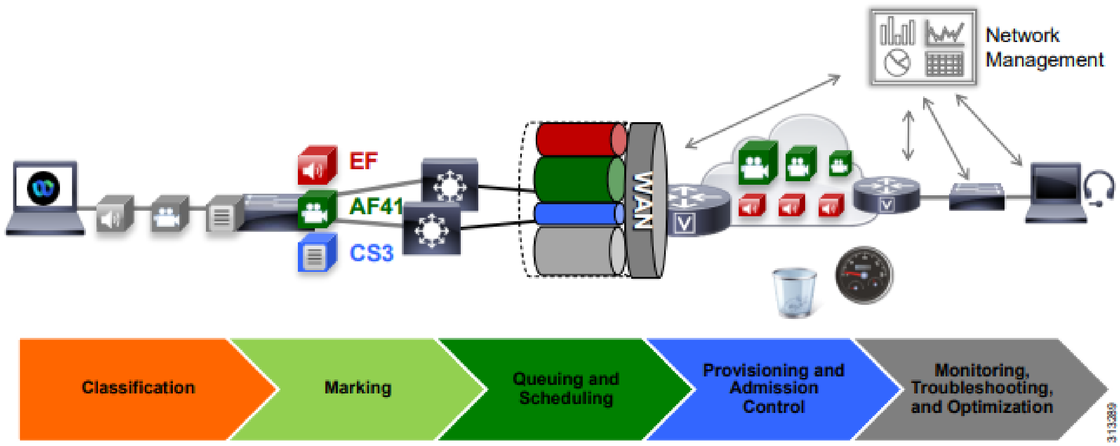


Figure 1.

Video Traffic: Bandwidth Usage High-definition Video Call

(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

Figure 1 Architecture for Bandwidth Management



Recommended Deployment

Modify the existing on-premises QoS switch and WAN and Internet policies to include Webex Services identification, classification, and marking.

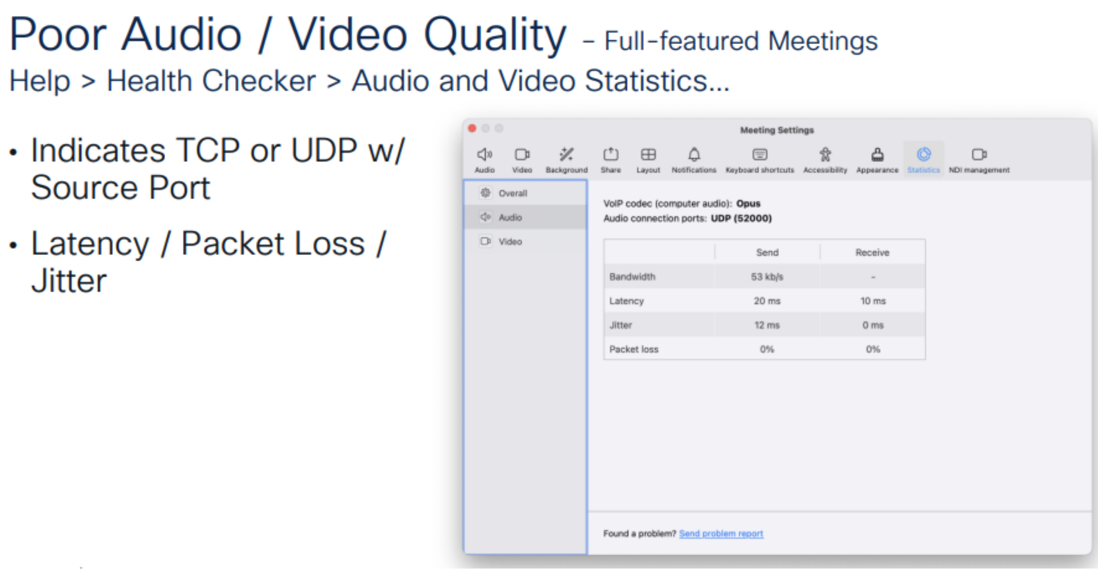
- Identify and classify media and signaling traffic from the Webex App and associated workloads as well as Webex Devices.
- Media and signaling marking recommendations:
  1. Mark all audio with Expedited Forwarding class EF (includes all audio of voice-only calls as well as audio for all types of video calls).
  2. Mark all Webex App video with an Assured Forwarding class of AF41 for a prioritized video class of service and AF42 for an opportunistic class of service. The marking of AF41 or AF42 will depend on the choice of whether to deploy opportunistic video during the on-premises deployment phase.
  3. Mark all call signaling with CS3. (All call signaling in HTTPS traffic will be marked based on the enterprise's current policy of traffic marking for HTTP/HTTPS – unless using NBAR which is covered in detail below)
- Configure QoS on all media originating and terminating applications such as the Video Mesh Nodes, Expressway and Cisco Unified Border Element.
- Update the WAN edge ingress re-marking policy.
- Update the WAN edge egress queuing and scheduling policy.

Every Cisco video endpoint employs several smart media techniques to avoid network congestion, recover from packet loss, and optimize network resources. The following smart media techniques are just some of the techniques employed by Cisco Webex App and Devices:

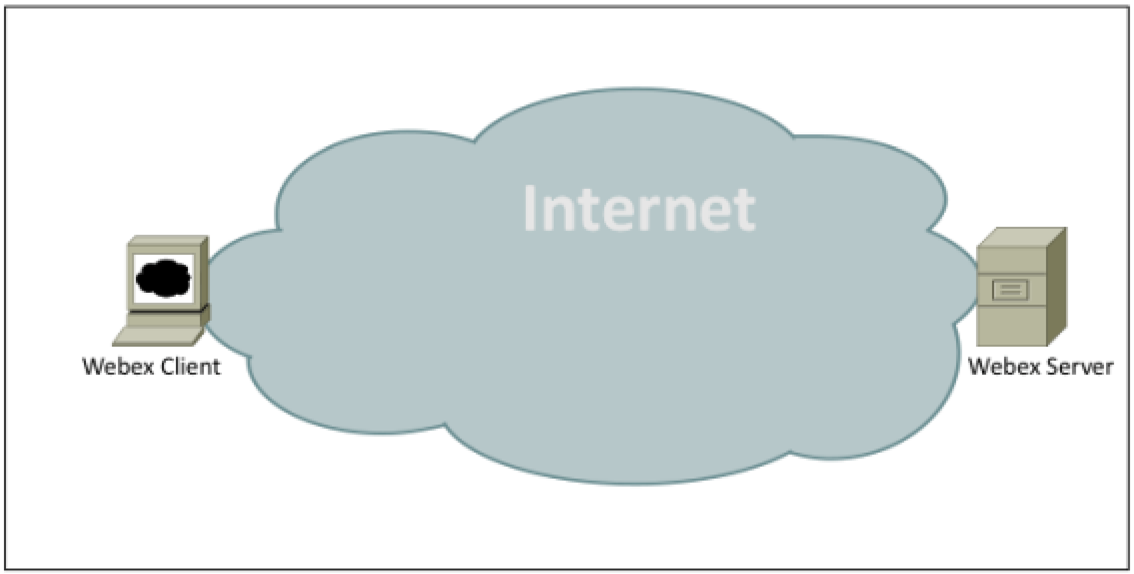
Media resilience techniques

- Encoder pacing
- Forward Error Correction (FEC)
- Rate adaptation

(https://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Collaboration/AltDesigns/BWM-Wbx.pdf.)



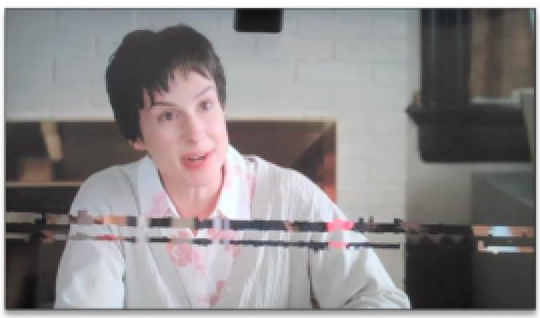
(https://www.ciscolive.com/c/dam/r/ciscolive/global-event/docs/2024/pdf/BRKCOL-3431.pdf.)

	There are three main factors that impact the quality of an audio or video call. These factors are **packet loss**, **latency**, and **jitter**. As shown in Figure 2, **packet loss is simply losing one or more packets within a stream of packets**. In this example, a Voice over IP (VoIP) packet is lost going from the Webex client to the Webex server. Loss can occur for a number of reasons in a network but most commonly it is caused by congestion or resource contention in the network or on the endpoints. This leads to routers, switches, or the endpoints themselves dropping or delaying packets. In other cases, packets encounter such a high delay that they are received too late to be played out. These packets will also be counted as lost.  ***Figure 2:*** *Packet Loss Example*

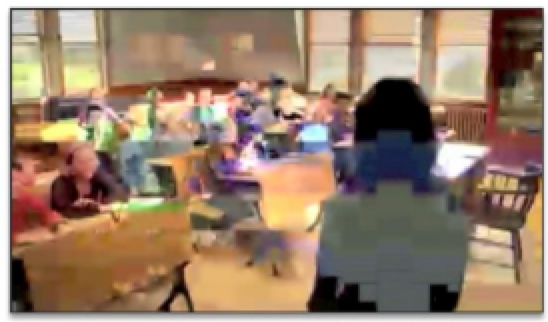
***Note**: Control Hub Diagnostics highlights potential problems with packet loss and delay by coloring voice and video streams as Good, Fair, or Poor. It is important to understand that just because Control Hub Diagnostics flags part of a conversation as Fair or Poor, it does not necessarily mean that the user had a bad experience. Voice and video compensation mechanisms and algorithms in combination with other factors can mitigate the impact of loss and delay to the point where it is not noticeable by the user.*

Video Quality Artifacts

The first step in troubleshooting video quality is to clearly understand the symptom experienced by the end user. Getting a screenshot or a picture of the display screen usually gives more information than a verbal description of the problem. The type of video artifacts observed in a screenshot or picture can be used to determine the troubleshooting path. For example, in Figure 13, pictures (a) and (b) are artifacts that are caused when video streams experience packet loss. Just by seeing the picture, you can immediately start troubleshooting to identify the source of the packet loss.



a. Video with line stripe artifacts



b. Frozen video with block artifacts

<table>
<tr>
<td></td>
<td>
Webex Control Hub makes it quite easy to find the device type and operating system version used by the end user to join a Webex meeting. All you need is the end user's email address and meeting time to find the right meeting and to get end user's device details. Figure 14 shows the list of meetings attended by a participant with the email address, rtpmsuser1@gmail.com.

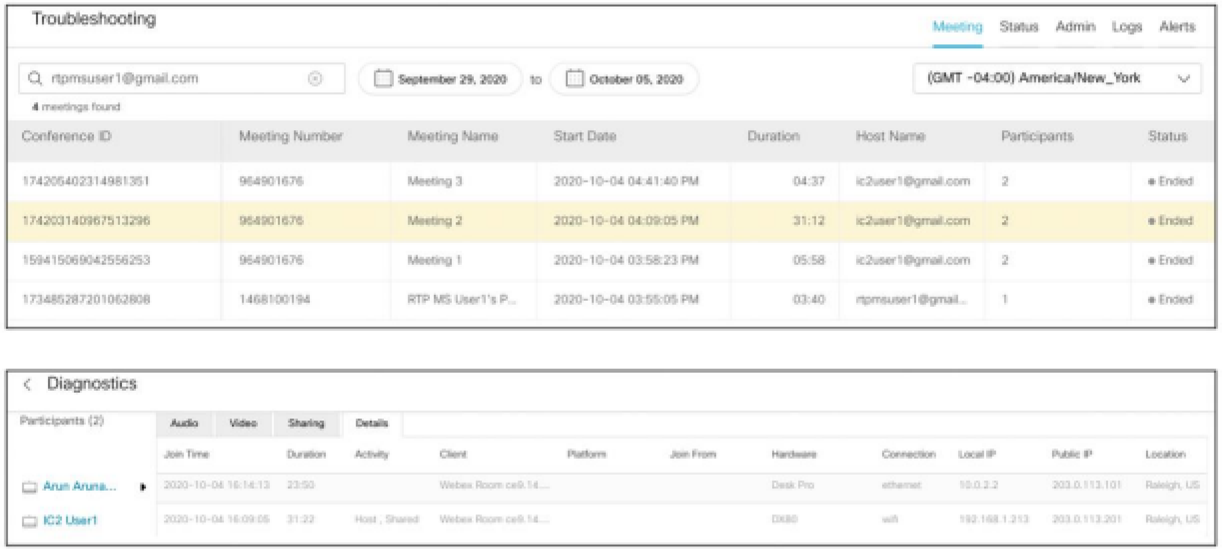


***Figure 14**: Participant Device Details*

(https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)
</td>
</tr>
<tr>
<td>forwarding the base layer to at least two of the</td>
<td>The Accused Instrumentalities perform a method for transmitting video signals, comprising forwarding the base layer to at least two of the plurality of video receivers via the internet protocol network; and
selectively forwarding one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions.</td>
</tr>
</table>

plurality of video receivers via the internet protocol network and selectively forwarding one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-	For example, because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer. **G.8.8.2 Specification of the base layer bitstream** Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. **Rec. ITU-T H.264 (06/2019)** 595 If a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth-limiting condition, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For instance, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.

limited conditions;	The sub-bitstream is derived by applying the following operations in sequential order: 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 489.

The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation.

See ITU-T H.264 at 624.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

	For example, the Accused Instrumentalities selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the plurality of video receivers through the internet protocol network based upon the identified bandwidth-limited conditions:

Table 4. Webex Meetings Bandwidth per Resolution Table

Layer	Bandwidth Range
90p active thumbnail (each)	~60-100 kb/s
180p main video	125-200 kb/s
360p main video	470-640 kb/s
720p main video	900k-1.5 mb/s
Content sharing (sharpness, 1080p/5)	120k – 1.3 mb/s
Content sharing (motion, 720p/30)	900k – 2.5 mb/s

Webex Meetings Desktop App Bandwidth Controls

Webex administrators have 2 key controls to help control bandwidth as used by clients that connect to Webex meetings should they choose to. Namely, you can cap the meeting layouts at either 360p as the max available resolution, or to enable 720p layers. Whether your site is administered on Webex Control Hub or Webex Site Administrator, the following controls are available in Configuration > Common Site Settings > Options:

☐ Turn on high-quality video (360p) *(Meetings, Training, Events and Support)*

☐ Turn on high-definition video (720p) *(Meetings, Training and Events)*

Figure 5.

Webex Meetings Desktop App Bandwidth Controls

Webex Media Improvements

The following are media improvements that have occurred in releases from 40.7 – 40.10

Improved Video Experience over WIFI (Defer video down-speeding): WIFI environments present a specific set of challenges and through testing we noticed the algorithms were too quick to drop resolution/bandwidth in certain WIFI environments. In 40.7 the algorithms have been updated to 'defer the down-speeding" of video for brief WIFI specific network dropouts. This resulted in improved user experience over WIFI. See figure5 for an illustration of this. When latency spikes for a specific duration the improved algorithms can detect that this is a WIFI specific event and not thus delay the down-speeding of the video resolution until that duration (n) is exceeded in which case would infer a different network event increasing latency.

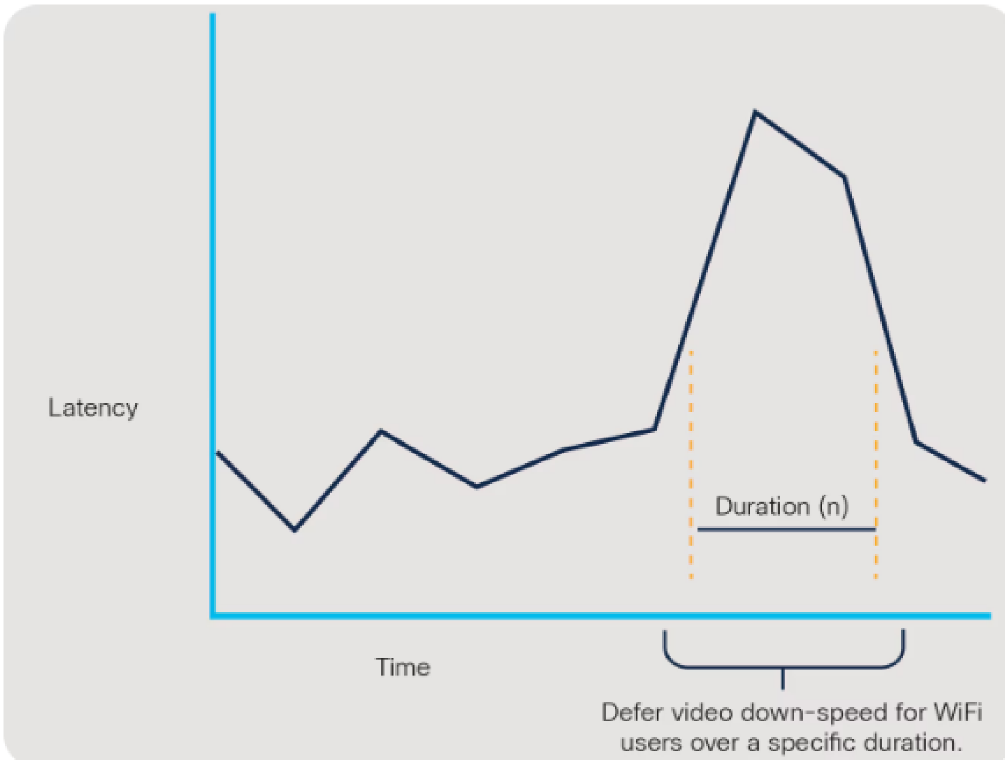


Figure 6.

Deferred Video Down-speeding

	Video Super Scaling is a technique used for users on sub-optimal networks. We provide 720p like experiences to users who are experiencing network impairments resulting in 360p connections to our servers for their meeting. Our updated scaling algorithm provides enhanced video quality when this degradation occurs, making network impairment downgrades less quality impacting. If a 360p image is simply scaled geometrically to a 720p resolution, the quality has not improved, but our technique involves migrating video post-processing tasks to the client for some video flows versus doing it in the cloud, resulting in being able to keep more precious bits of information to reconstruct the video at a given bandwidth. In a meeting as shown in figure 8 and 9, the bandwidth requirements are as follows. Assuming the receiver is not bandwidth restricted, Webex app will adapt to available network conditions. Table 8 outlines the bandwidth utilization. Figure 13 illustrates the Active Speaker layout which is the default meeting layout for Webex Teams App. **Table 8. Cisco Webex Teams Bandwidth Utilization**

Meeting Scenario	Main Video+Audio (no content sharing)	Main Video+Audio (content sharing)
Webex Teams 1:1 Call	1.5 mb/s peak, 1.3 mb/s avg	1.5 mb/s peak, 1.3 mb/s avg
Webex Teams Meeting (fig 13 layout), 6 participants	2.3 mb/s peak, 1.7 mb/s avg	2.9/ mb/s peak, 2.0 mb/s avg
Webex Teams Meeting (fig 14 layout), 6 participants	2.5 mb/s peak, 1.8 mb/s avg	3.5 mb/s peak, 2.3 mb/s avg

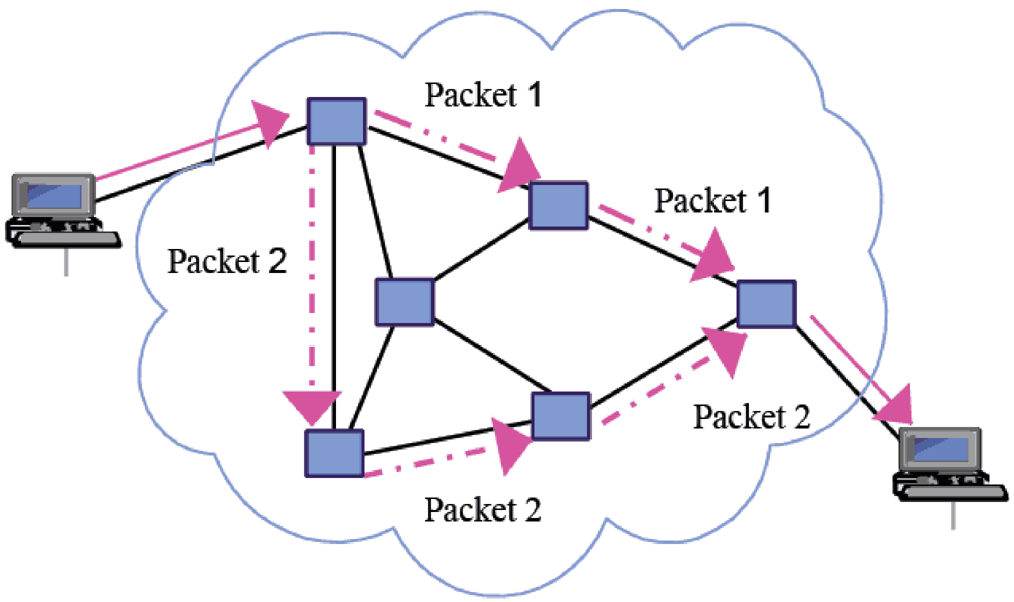
(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets,	The Accused Instrumentalities perform a method for transmitting video signals, wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets. For example, scalable video routers facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet, using internet protocol. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.

networkencyclopedia.com/packet-switching/

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or "packets" that contain sequence numbers so that they can be reassembled at the destination.



Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

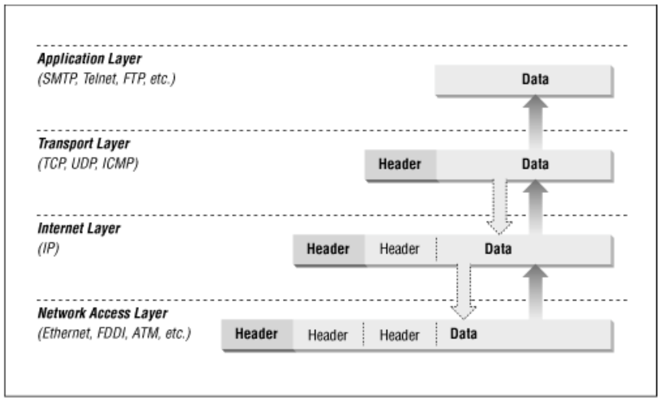
Not Secure | web.deu.edu.tr/doc/oreily/networking/firewall/ch06_03.htm

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. Figure 6.2 shows how this works.

Figure 6.2: Data encapsulation

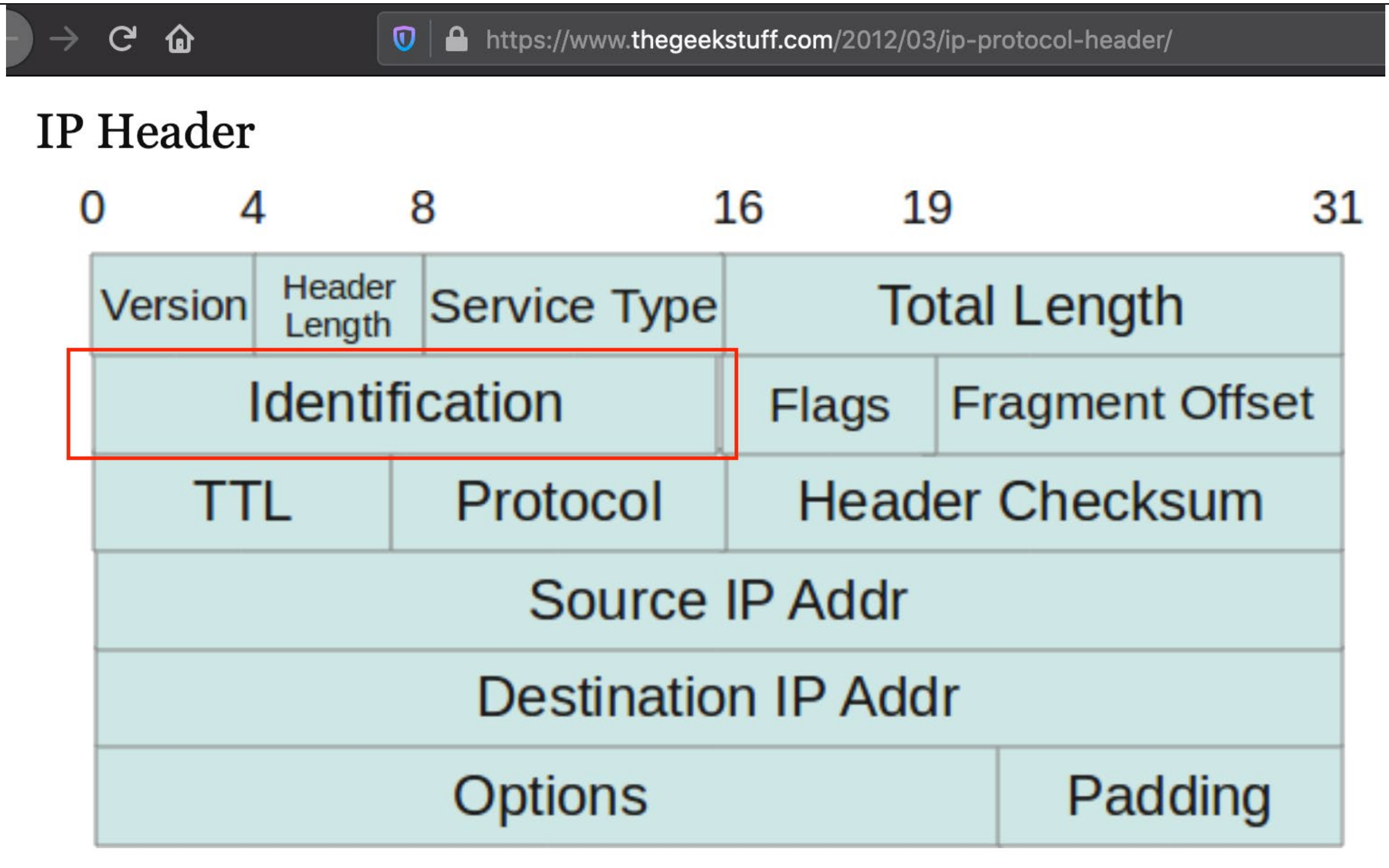


At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

	ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf *A. Network Abstraction Layer (NAL)* The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-
each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.	The Accused Instrumentalities perform a method for transmitting video signals, wherein the layered video data stream is transmitted according to an internet protocol; and wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. For example, each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence:

	Not Secure | linfo.org/packet_header.html **Packet Header Definition** A *packet header* is the portion of an IP (Internet protocol) *packet* that precedes its body and contains addressing and other data that is required for it to reach its intended destination. Packets are the fundamental unit of information transport in all modern computer networks, and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the *payload*, and a *trailer*. The header's format is specified in the Internet protocol. It normally contains 20 bytes of data, although an option exists within it that allows the addition of more bytes. Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's IP address, the intended receiver's IP address, the number of packets the message has been broken into, the identification number of the particular packet, the protocol (e.g. 1 for ICMP, 2 for IGMP, 6 for TCP and 17 for UDP) used, the packet length (on networks that have variable length packets), the *time to live* (i.e., the number of links or *hops* that the packet can be routed before being allowed to expire) and synchronization data (several bits that help the packet match up to the network).


https://www.thegeekstuff.com/2012/03/ip-protocol-header/
IP Header
0
4
8
16
19
31
Version
Header Length
Service Type
Total Length
Identification
Flags
Fragment Offset
TTL
Protocol
Header Checksum
Source IP Addr
Destination IP Addr
Options
Padding

https://www.thegeekstuff.com/2012/03/ip-protocol-header/

- ***Identification(16 bits)***: This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

For example, each data packet is encoded with a layer identifier, wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.

Each scalable layer is associated with a unique layer identifier as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier layerId does not include any scalable layer with a layer identifier greater than layerId, but it may include scalable layers with layer identifiers less than layerId. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause.

NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have sub_pic_layer_flag[i] equal to 0, the unique layer identifier values may be set equal to (128 * dependency_id + 8 * quality_id + temporal_id), with dependency_id, quality_id, and temporal_id being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer.

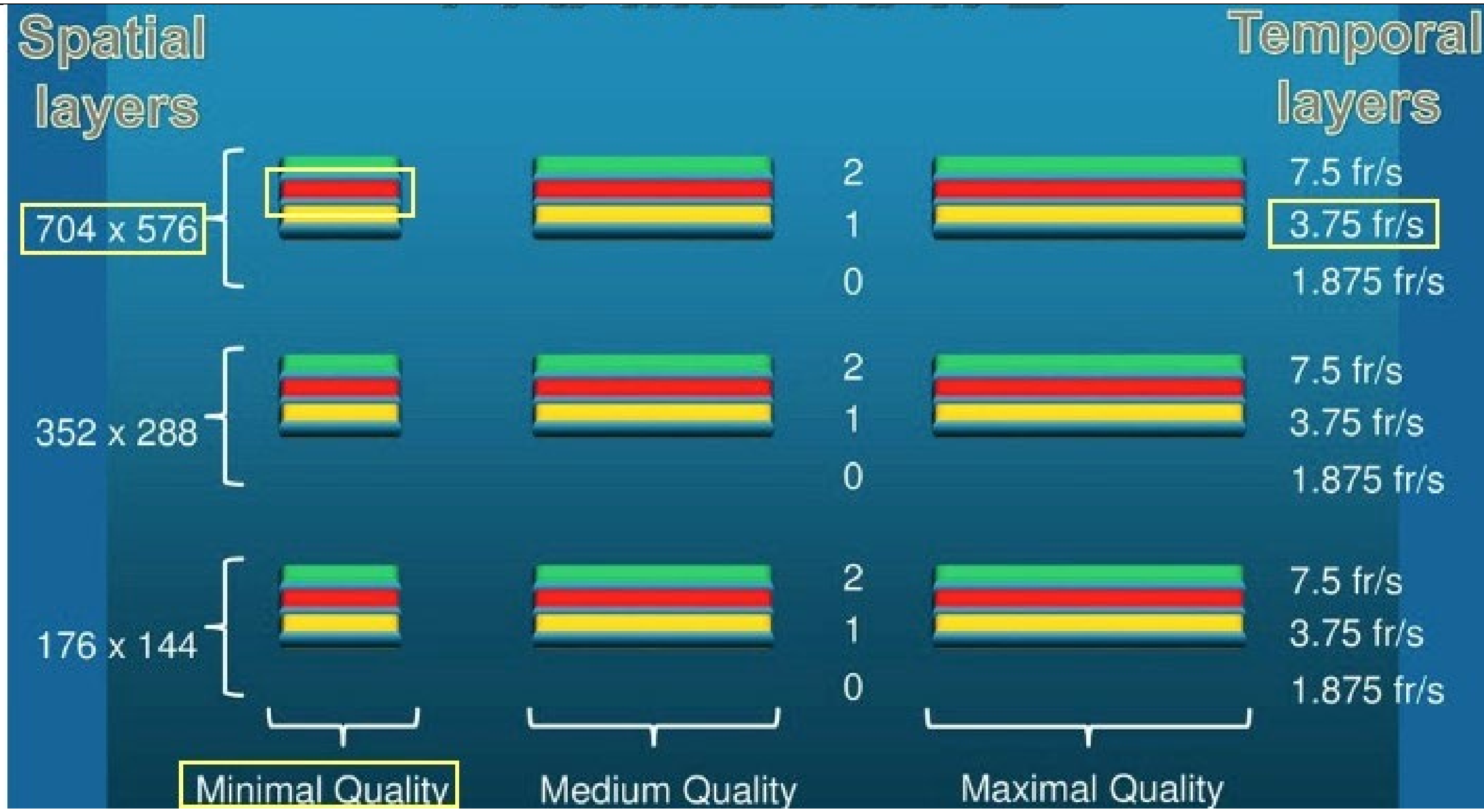
See ITU-T H.264 at 624.

NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including priority_id dependency_id, quality_id, and temporal_id, according to the standard.

dependency_id specifies a dependency identifier for the NAL unit. dependency_id shall be equal to 0 in prefix NAL units. The assignment of values to dependency_id is constrained by the sub-bitstream extraction process as specified in clause G.8.1.

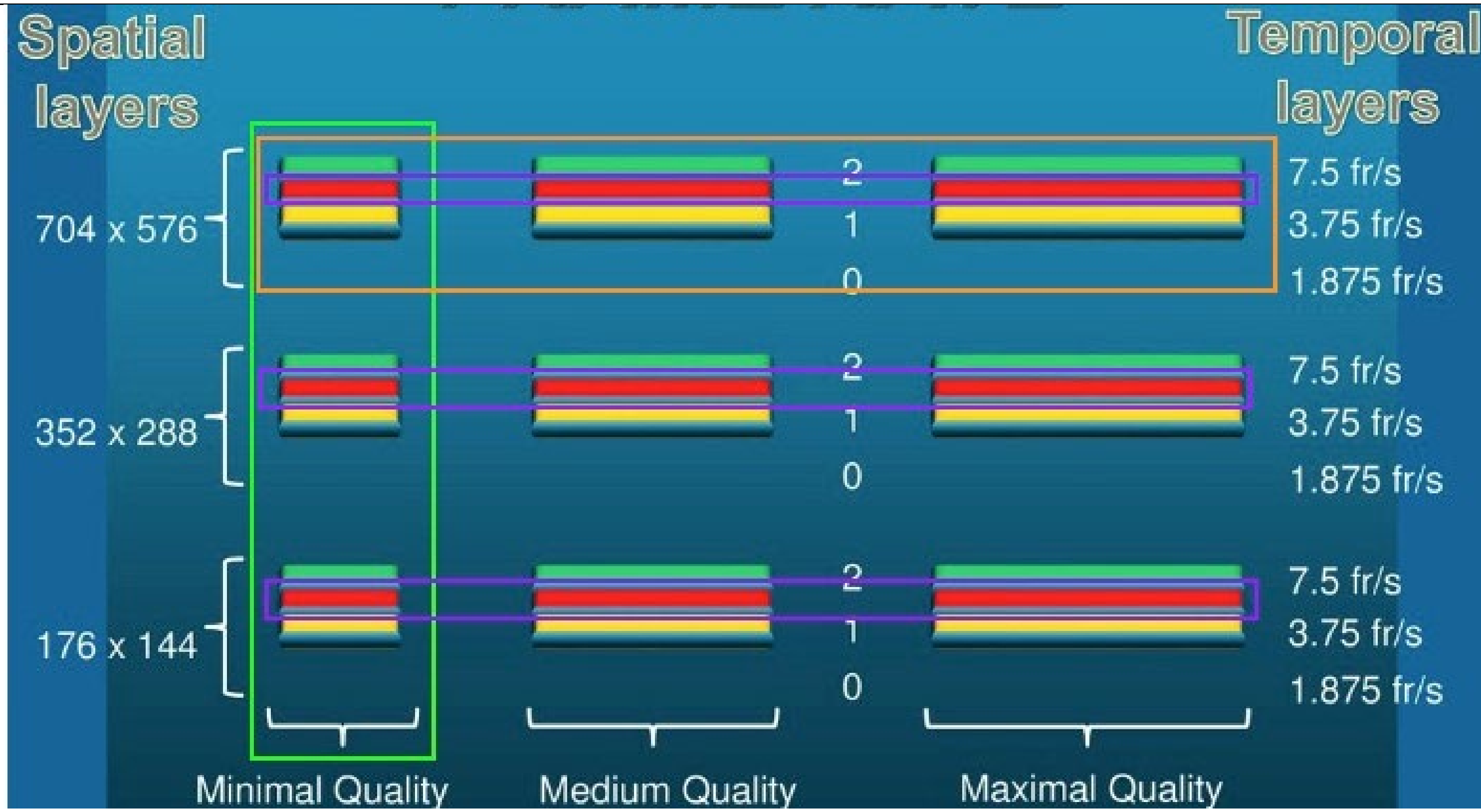
See ITU-T H.264 at 459.

	priority_id specifies a priority identifier for the NAL unit. The assignment of values to priority_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458. **quality_id** specifies a quality identifier for the NAL unit. quality_id shall be equal to 0 in prefix NAL units. The assignment of values to quality_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable DQId is derived by DQId = (dependency_id << 4) + quality_id (G-63) When nal_unit_type is equal to 20, the bitstream shall not contain data that result in DQId equal to 0. See ITU-T H.264 at 459. **temporal_id** specifies a temporal identifier for the NAL unit. The assignment of values to temporal_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The value of temporal_id shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with nal_unit_type equal to 5 or idr_flag equal to 1, temporal_id shall be equal to 0. See ITU-T H.264 at 459. For instance, the graphic below illustrates an example of a hierarchal virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.



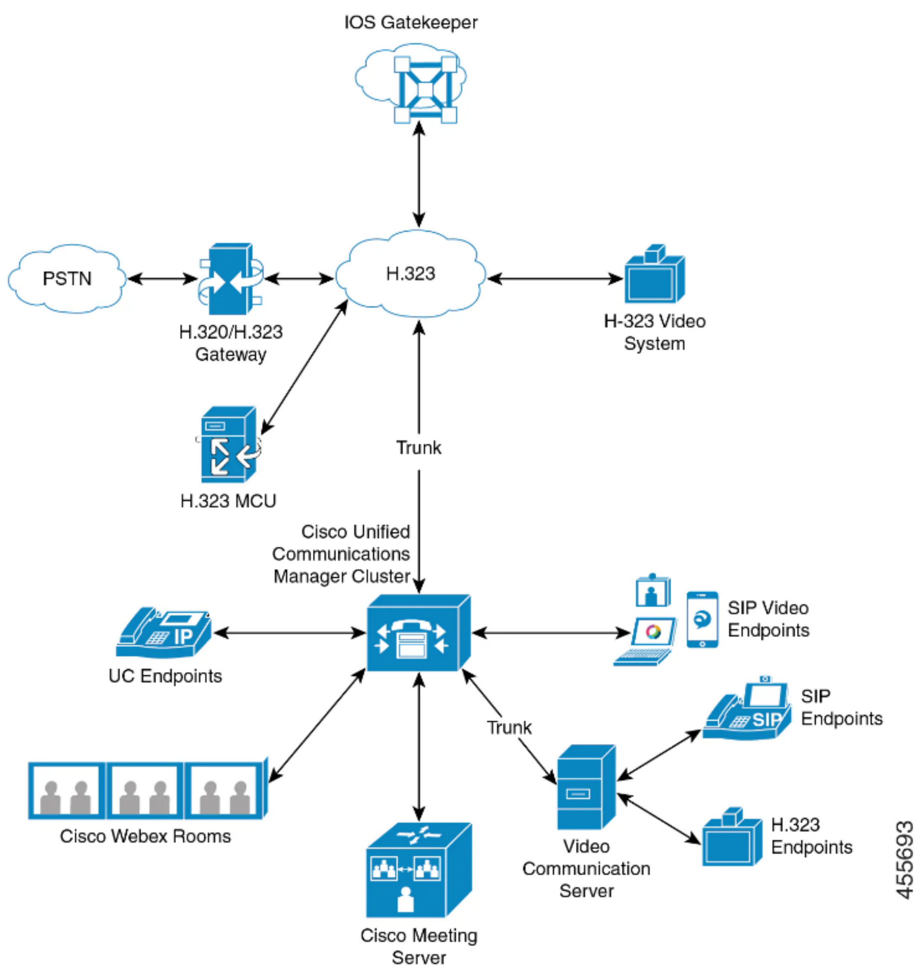
Source: https://slideplayer.com/slide/13389021/

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.

	 Source: https://slideplayer.com/slide/13389021/
11. A scalable video coding router, comprising:	The Accused Instrumentalities include a scalable video coding router. For example, scalable video routers conforming to the H.264 standard perform a method of transmitting video signals by forwarding IP data packets containing scalable video information.

Video Network

The following illustration provides an example of a video network that uses a single Unified Communications Manager cluster. In a successful video network, any endpoint can call any other endpoint. Video availability only exists if both endpoints are video-enabled. Video capabilities extend across trunks.



	The Cisco video conference portfolio comprises the following video bridges: • Cisco TelePresence MCU series • Webex Meeting Server (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Figure 11-14 Cisco WebEx Collaboration Cloud Architecture

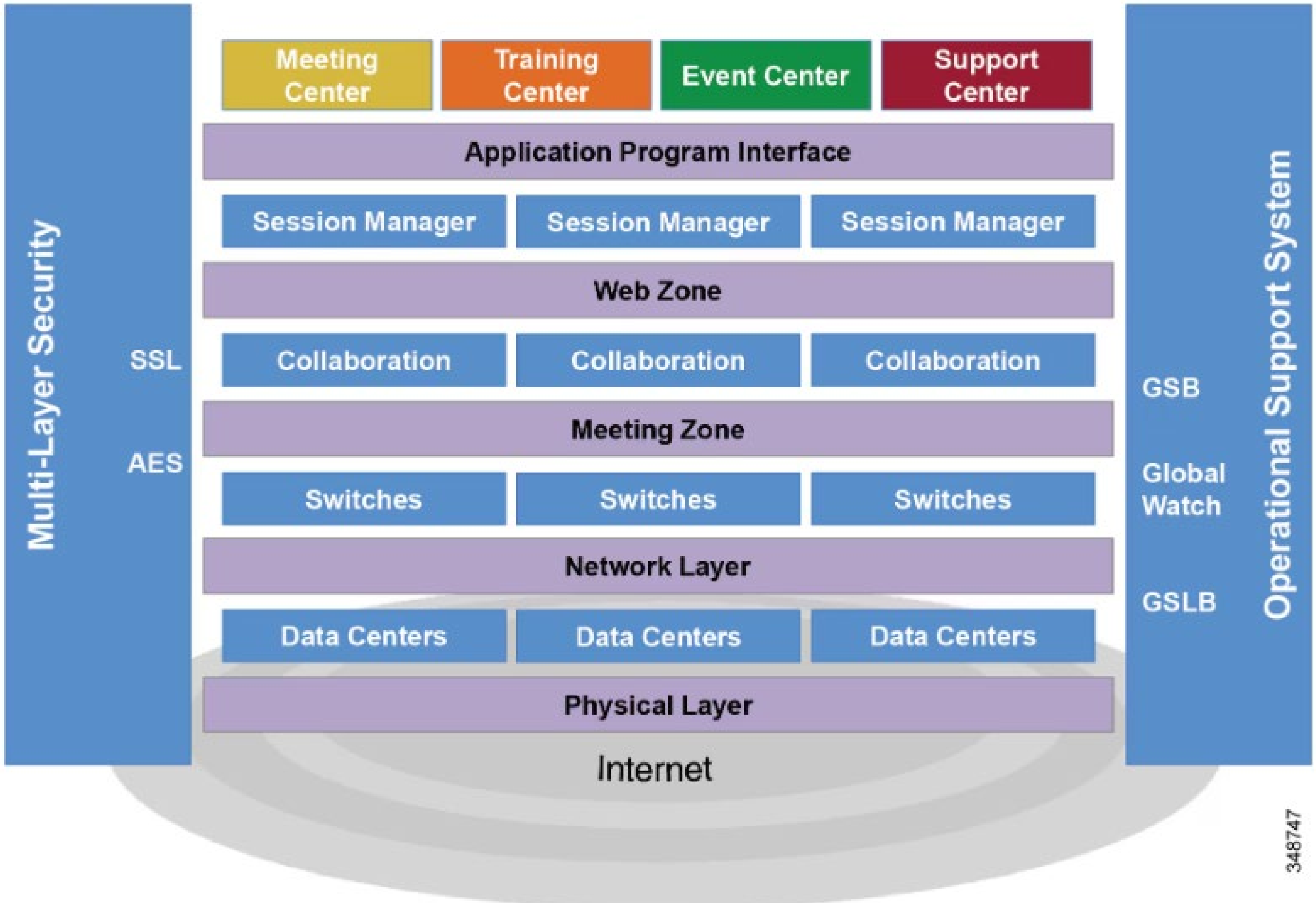
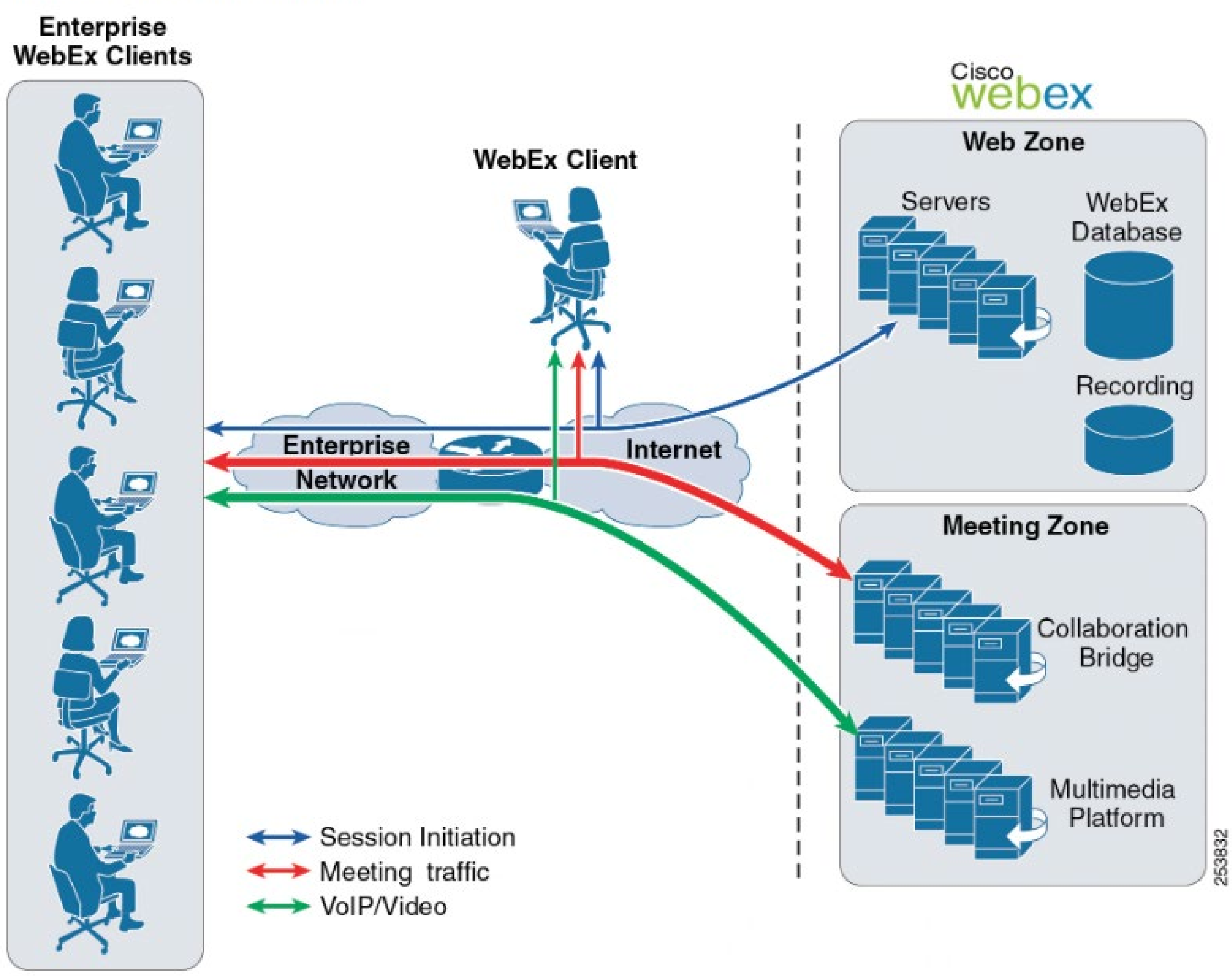


Figure 11-15 WebEx Deployment



	Cisco Rich Media Conferencing consists of the conferencing solutions described below. The details pertaining to each solution are described in each individual section that follows. • Cisco Unified CM Audio Conferencing This solution allows Unified CM to use its internal software component or external hardware digital signal processors (DSPs) as the resources to perform audio conferencing. • Cisco Meeting Server Cisco Meeting Server is an on-premises video conferencing solution. Each user has a personal Space that can be used to conduct meetings. Users can manage items such Space creation, adding members to a Space, and PIN creation from the Cisco Meeting App. • Cisco Collaboration Meeting Rooms Hybrid Cisco CMR Hybrid combines the on-premises video conference and the WebEx Meeting Center conference into a single meeting, which allows TelePresence and WebEx participants to join and share voice, video, and content. CMR Hybrid meetings can be either scheduled or non-scheduled. • Cisco WebEx Meeting Center Video Conferencing Cisco WebEx Meeting Center Video Conferencing (formerly Cisco Collaboration Meeting Rooms (CMR) Cloud) is an alternate conferencing deployment model that does not require any on-premises conferencing resources or management infrastructure. It supports both scheduled and non-scheduled meetings as well as TelePresence, audio, and WebEx participants in a single call, all hosted in the cloud. • Cisco WebEx Meetings Server Where cloud-based web and audio conferencing is not suitable, it is possible to use the on-premises WebEx Meetings Server solution. This product offers a standalone audio, video, and collaboration web conferencing platform. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)

	The Accused Instrumentalities implement the scalable video coding features in Annex G of the H.264 standard. For example: **Video Calls** The typical video call includes two or three Real-Time Protocol (RTP) streams in each direction (that is, four or six streams). The call can include the following stream types: • Video (H.261, H.263, H.263+, H.264-SVC, X-H.264UC, H.264-AVC, H.265, AV1 and VT Camera wideband video codecs) • Far-End Camera Control (FECC) - Optional • Binary Floor Control Protocols (BFCP) **SIP Video** SIP video supports the following video calls by using the SIP Signaling Interface (SSI): • SIP to SIP • SIP to H.323 • SIP to SCCP • SIP intercluster trunk • H.323 trunk • Combination of SIP and H.323 trunk SIP video calls also provide media control functions for video conferencing. Unified Communications Manager video supports SIP on both SIP trunks and lines support video signaling. SIP supports the H.261, H.263, H.263+, H.264 (AVC), H.264 (SVC), X-H.264UC (Lync), and AV1 video codecs (it does not support the wideband video codec that the VTA uses).

	Video Codecs Common video codecs include H.261, an older video codec, H.263, a newer codec that gets used to provide internet protocol (IP) video, and H.264, a high-quality codec. The system supports H.264 for calls that use the Skinny Client Control Protocol (SCCP), H.323, and SIP on originating and terminating endpoints only. The system also supports regions and locations. Unified Communications Manager maintains the offerer's video codec ordering preference when making the answer, if possible. H.265 is the preferred video codec is available on the endpoints, otherwise, Unified Communications Manager follows the following codec preference preference order:

Preference Order	Codecs	Description
1	H.265 (HEVC)	Provides higher quality video using lower bandwidth.
2	H.264 (SVC)	Allows rendering of variable quality video from the same media stream, by disregarding a subset of the packets received. **Note** H. 264 SVC is a new annex to the H.264-AVC video compression standard; meaning it is an enhancement on top of H.264-AVC. It provides the ability to encapsulate multiple video streams at various frame-rates and resolutions in one container.

(https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

	Meeting Center uses the H.264 AVC/SVC codec to provide high-definition video for the conference. Higher network bandwidth is needed for those deployments. For further details regarding network traffic optimization for high-definition video, see Capacity Planning. **Network Traffic Planning** Network traffic planning for Cisco WebEx Meeting Center Video Conferencing consists of the following elements: • WebEx Clients bandwidth The WebEx meeting client uses Scalable Video Coding (SVC) technology to send and receive video. It uses multi-layer frames to send video, and the receiving client automatically selects the best possible resolution to receive video that typically requires 1.2 to 3 Mbps of available bandwidth. For more information regarding network traffic planning for WebEx clients, refer to the *Cisco WebEx Network Bandwidth* white paper, available at https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meeting-center/white_paper_c11-691351.html • Bandwidth for video device from enterprise to WebEx Cloud For optimal SIP audio and video quality, Cisco recommends setting up the video bandwidth for at least 1.5 Mbps per device screen in the region associated with the endpoint registering with Cisco Unified CM. For example, if a triple-screen device registers with Unified CM, video bandwidth of 4.5 Mbps should be allocated in the associated region. (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)
a memory; and a processor,	The Accused Instrumentalities include a scalable video coding router, comprising a memory and a processor, wherein the processor executes instructions stored in the memory to cause the scalable video coding router to perform the claimed steps.

wherein the processor executes instructions stored in the memory to cause the scalable video coding router to:	For example, all routers comprise a memory and a processor, wherein the processor executes instructions stored in the memory:

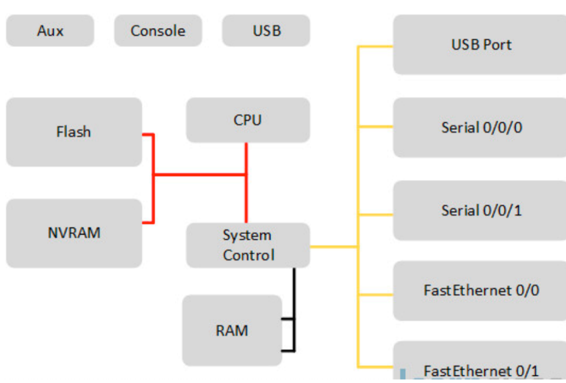
https://www.learncisco.net/courses/icnd-2/network-management/router-internal-components-overview.html

Router Internal Components Overview

I want you to take a moment and look at the components we have listed here.

- CPU - Executes operating system instructions
- RAM (Random access memory) - Contains the running copy of the configuration file. Stores the rougint table. RAM contents lost when power is off.
- ROM (Read-only memory) - Holds diagnostic software used when router is powered up. Stores the router's bootstrap program.
- NVRAM (Nonvolatile RAM) - Stores startup configuration. This may include IP addresses (routing protocol, hostname of a router).
- Flash memory - Contains the operating system (Cisco IOS)
- Interfaces - Multiple fisical interfaces that are used to connect the network. Fast Ethernet, Gigabit Ethernet; Serial; Management interfaces

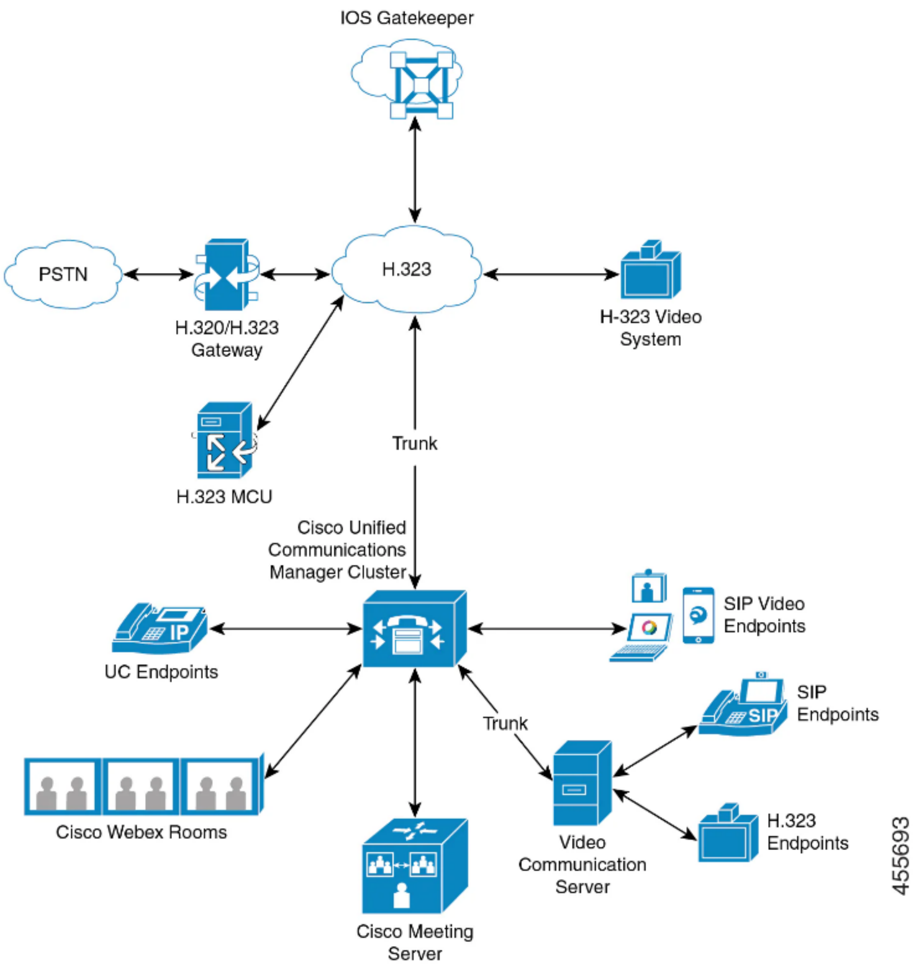
CPU, RAM, ROM, interfaces - those all sound very familiar to something else. A computer, a PC! All of these components we find inside the tower of our PC or inside of our laptop. So we have our router, we look inside, we have the same types of components. Let's go through these individually.



	For example:

Video Network

The following illustration provides an example of a video network that uses a single Unified Communications Manager cluster. In a successful video network, any endpoint can call any other endpoint. Video availability only exists if both endpoints are video-enabled. Video capabilities extend across trunks.



	The Cisco video conference portfolio comprises the following video bridges: • Cisco TelePresence MCU series • Webex Meeting Server (https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/admin/14SU2/cucm_b_feature-configuration-guide-for-cisco14su2/cucm_m_video-telephony.html.)

Figure 11-14 Cisco WebEx Collaboration Cloud Architecture

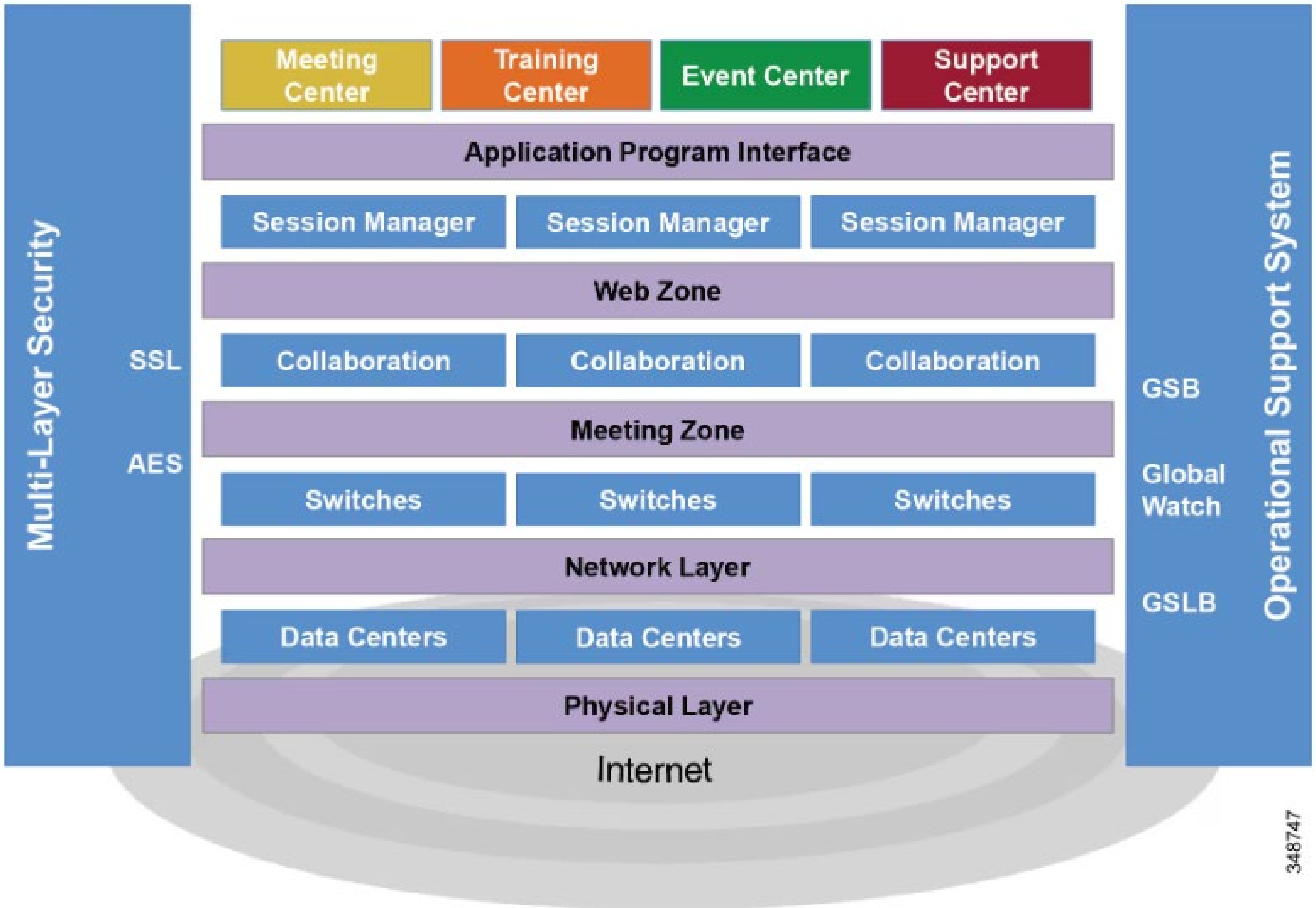
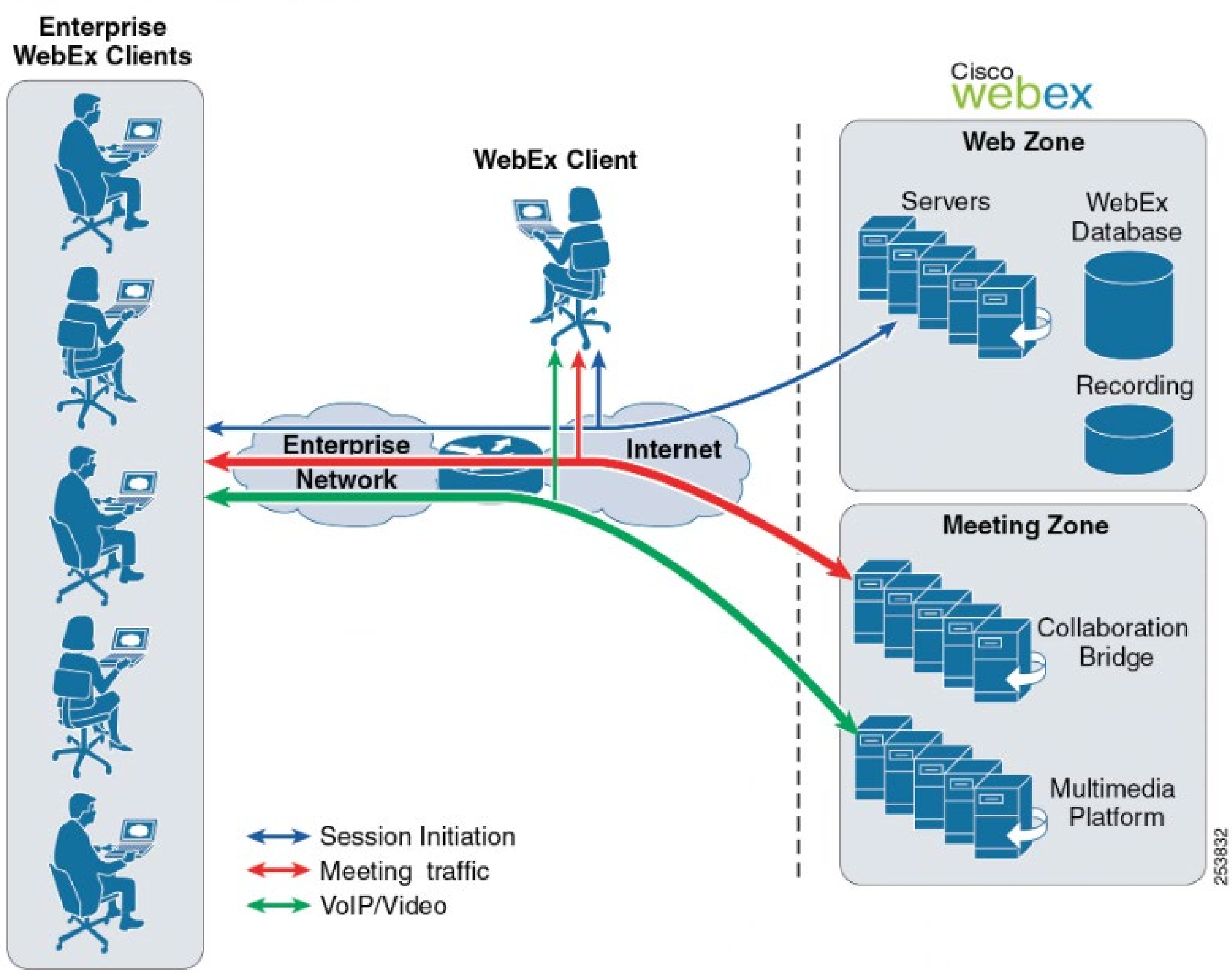


Figure 11-15 WebEx Deployment



(https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/cucm/srnd/collab12/collab12/confernc.html.)

Product Overview

Cisco Webex Meetings Server is a virtualized, software-based solution that runs on Cisco Unified Computing System™ (Cisco UCS®) servers and VMware. It uses virtual appliance technology for rapid turn-up of services to end users. With Cisco Webex Meetings Server, there are two options for enabling mobile users to more securely access Cisco Webex conferences without going through a VPN. The first option is to deploy reverse proxy (or edge servers) in the enterprise perimeter (or DMZ). The second option, shown in Figure 1, is to deploy the reverse proxy servers behind your internal firewall, thus eliminating all DMZ components and related information security concerns.

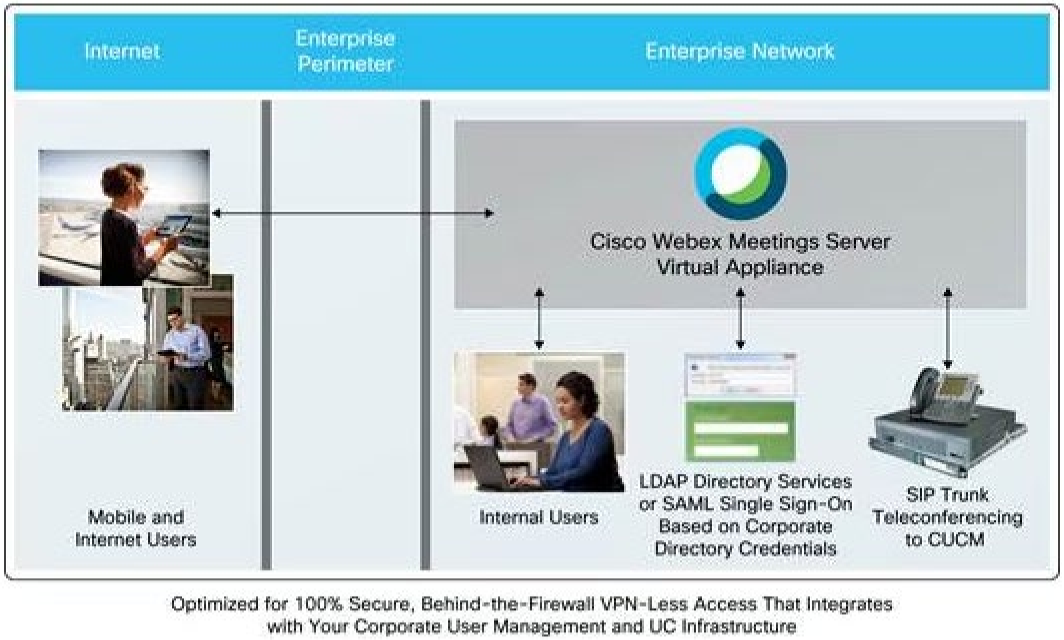


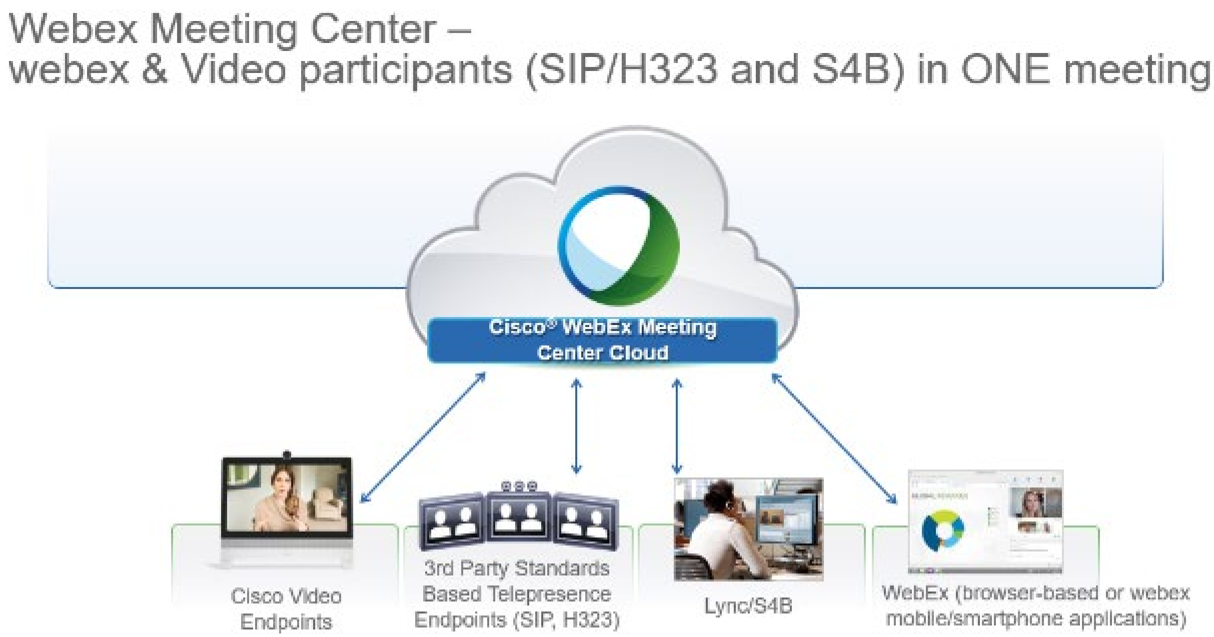
Figure 1.

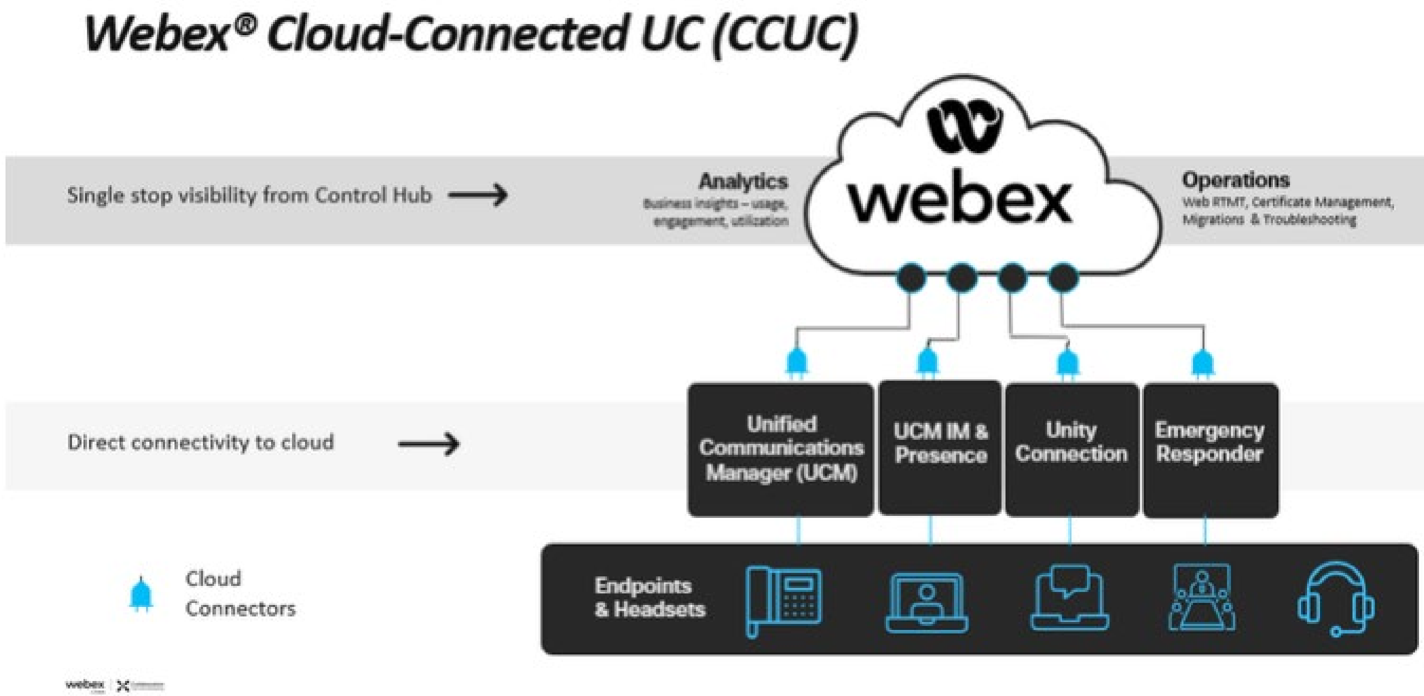
Full Deployment of Cisco Webex Meetings Server Behind a Firewall

System Requirements

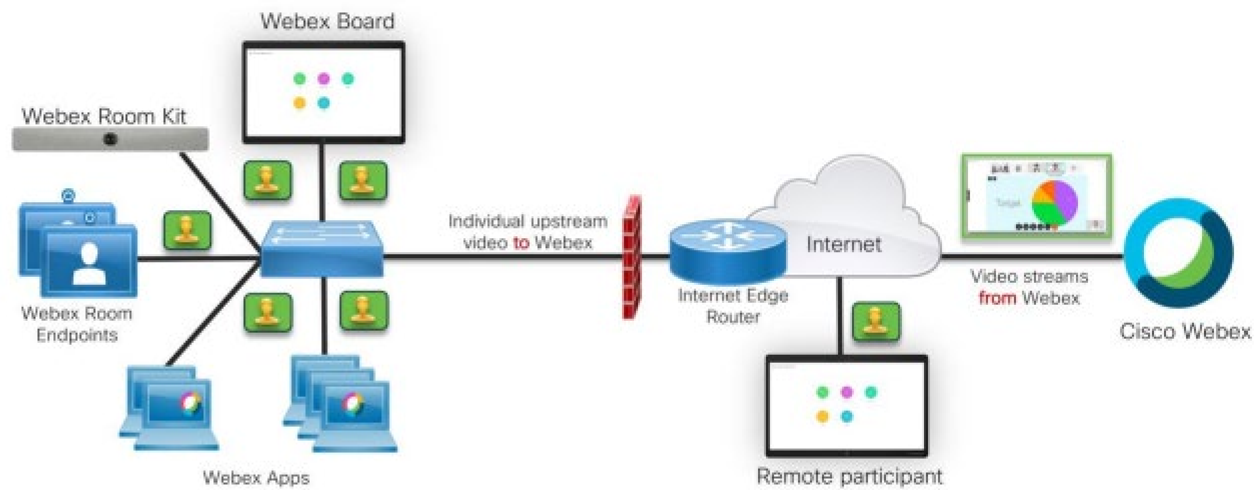
Cisco Webex Meetings Server is compatible with Cisco UCS servers that meet or exceed the specifications presented in this section.

Module	Requirements
Host server	Cisco UCS C-Series rack server or equivalent B-Series blade server
Network interfaces	• Minimum 1 physical Network Interface Card (NIC) for a nonredundant configuration • Redundant configurations must have all NIC interfaces duplicated and connected to an independent switching fabric
Internal storage (direct attached storage [DAS]) for ESXi hosts where internal machines are deployed	Minimum of 4 drives in a RAID-10 or RAID-5 configuration
Internal storage (DAS)for ESXi hosts where Internet Reverse Proxy (IRP) virtual machines are deployed	Minimum of 2 drives in a RAID-1 configuration
SAN storage	Can be used as a substitute for DAS
Network-Attached Storage (NAS)	Can be used as a substitute for DAS or SAN
Hypervisor	• ESXi versions and vSphere licenses • 1 VMware license per processor socket
Email server	• Fully Qualified Domain Name (FQDN) of the mail server that the system uses to send emails • Port number: Default value of the Simple Mail Transfer Protocol (SMTP) port number is 25 or 465 (secure SMTP port number)

	(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings-server/datasheet-c78-717754.html.)  (https://community.cisco.com/t5/collaboration-knowledge-base/cisco-webex-meetings-overview/ta-p/3648888.)

	 (https://community.cisco.com/t5/collaboration-knowledge-base/webex-cloud-connected-uc-connect-your-cucm-to-webex-cloud-for/ta-p/4172103.)
receive a layered video data stream that comprises a base layer and a set of	The Accused Instrumentalities receive a layered video data stream that comprises a base layer and a set of enhancement layers. For example, the Accused Instrumentalities conform to the H.264 standard, including Annex G (SVC). Scalable video routers conforming to the H.264 standard perform a method of receiving a multi-layered stream.

enhancement layers,	sciencedirect.com/topics/engineering/scalable-video-coding **Communicating Pictures: Delivery Across Networks** David R. Bull, in Communicating Pictures, 2014 **Scalable video encoding** It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

	 ***Figure 11**: Video Streams in a Webex Meeting* (https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)

https://www.mistralsolutions.com/articles/h-264-svc-update-h-264-video-compression-standard/

+91 80 45621100 info@mistralsolutions.com

About us ▾ Products ▾ Services ▾

MISTRAL®
... Partners in Real Time

AEROSPACE AND DEFENSE
Land, Naval and Airborne

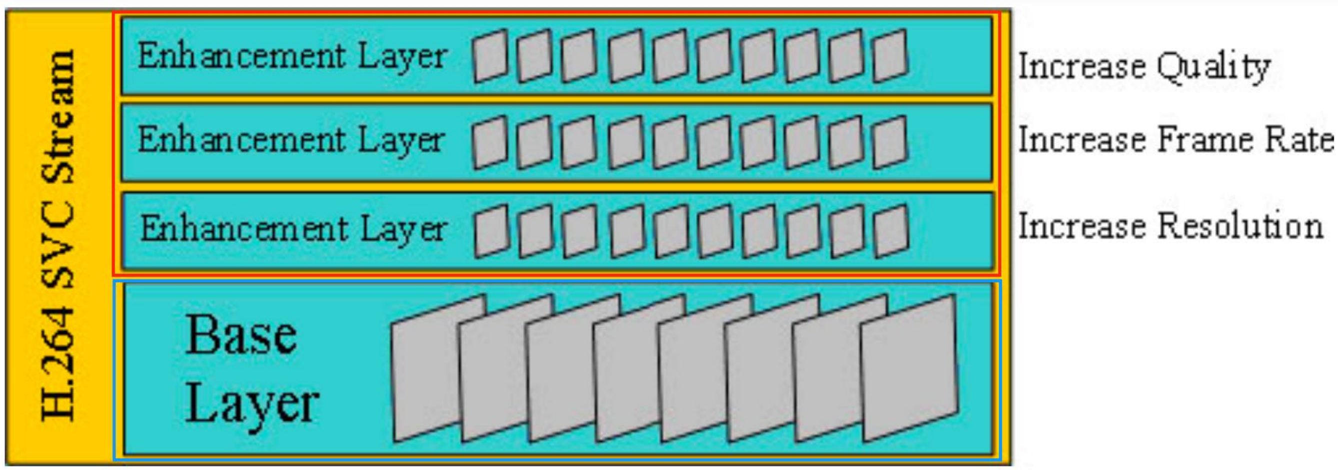


Figure 1: H.264 SVC Layering Technique

The "base layer" is the minimum decodable bitstream subset (or sub-bitstream), and every other layer is an enhancement layer.

G.3.4 **base layer**: A *bitstream subset* that contains all *NAL units* with the nal_unit_type *syntax element* equal to 1 and 5 of the *bitstream* and does not contain any *NAL unit* with the nal_unit_type *syntax element* equal to 14, 15, or 20 and conforms to one or more of the profiles specified in Annex A.

See ITU-T H.264 at 437.

https://www.sciencedirect.com/topics/engineering/enhancement-layer

MPEG-4 Visual and H.264/AVC: Standards for Modern Digital Video

Berna Erol, ... Lowell Winger, in The Essential Guide to Video Processing, 2009

10.3.6 Scalability

In addition to the video coding tools discussed so far, MPEG-4 Part 2 provides scalability tools that allow organization of the bitstream into base and enhancement layers. The enhancement layers are transmitted and decoded depending on the bit rate, display resolution, network throughput, and decoder complexity constraints. Temporal, spatial, quality, complexity, and object-based

	Scalable video data streams coded according to the H.264 standard comprise a single bitstream that is further comprised of layers, wherein combinations of layers form "bitstream subsets" that may be decoded as video data streams. **G.3.54 scalable bitstream**: A *bitstream* with the property that one or more *bitstream subsets* that are not identical to the scalable bitstream form another *bitstream* that conforms to this specification. See Telecommunication Standardization Sector, International Telecommunication Union, H.264 Infrastructure of audiovisual services – Coding of moving video, 437 (Jun. 2019) ("ITU-T H.264"). **G.3.6 bitstream subset**: A *bitstream* that is derived as a *subset* from a *bitstream* by discarding zero or more *NAL units*. A *bitstream subset* is also referred to as *sub-bitstream*. See ITU-T H.264 at 437. Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding. Otherwise, the coding algorithm as described in the previous paragraph is used. See ITU-T H.264 at 4.
identify bandwidth-limited conditions of an internet	The Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between the video router and a set of video receivers. For example, scalable video routers detect conditions in which bandwidth is limited. Such a bandwidth limitation could be caused by a congested state of a network over which packets must be sent, or a weak physical connection between

protocol network between the video router and a set of video receivers;	two or more links in the network, a limitation in the capabilities of a given receiver, or some other network condition that would limit the bandwidth between one or more nodes. sciencedirect.com/topics/engineering/scalable-video-coding **Communicating Pictures: Delivery Across Networks** David R. Bull, in Communicating Pictures, 2014 **Scalable video encoding** It is common in video transmission for the signal to be encoded and transmitted without explicit knowledge of the downstream network conditions. In cases where network congestion exists in a packet switched network such as the internet, routers will discard packets that they are unable to forward due to congestion. In some cases it is possible for packets within a stream to be prioritized or embedded such that the least important information is discarded before the more important data. This mechanism lends itself to scalable or layered encoding, where a video signal is composed of a hierarchical layering in terms of spatial resolution, temporal resolution or SNR. This enables devices to transmit and receive multilayered video streams where a base level of quality can be improved through the use of optional additional layers that enhance resolution, frame rate, and/or quality.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

For example, the Accused Instrumentalities identify bandwidth-limited conditions of an internet protocol network between the video router and a plurality of video receivers as shown below:

Overview

In this document we will discuss bandwidth utilization. Bandwidth values used will be in payload bit rate which does not include packetization overhead and are covered in 3 categories, average, peak and maximum bit rate:

Average (**avg**) is the average over time for a meeting participant.

Peak (**peak**) is the typical peak bursts over the same time period for a meeting participant.

Maximum (**max**) is the maximum bit rate that the device is capable of either due to device limitations or device configuration.

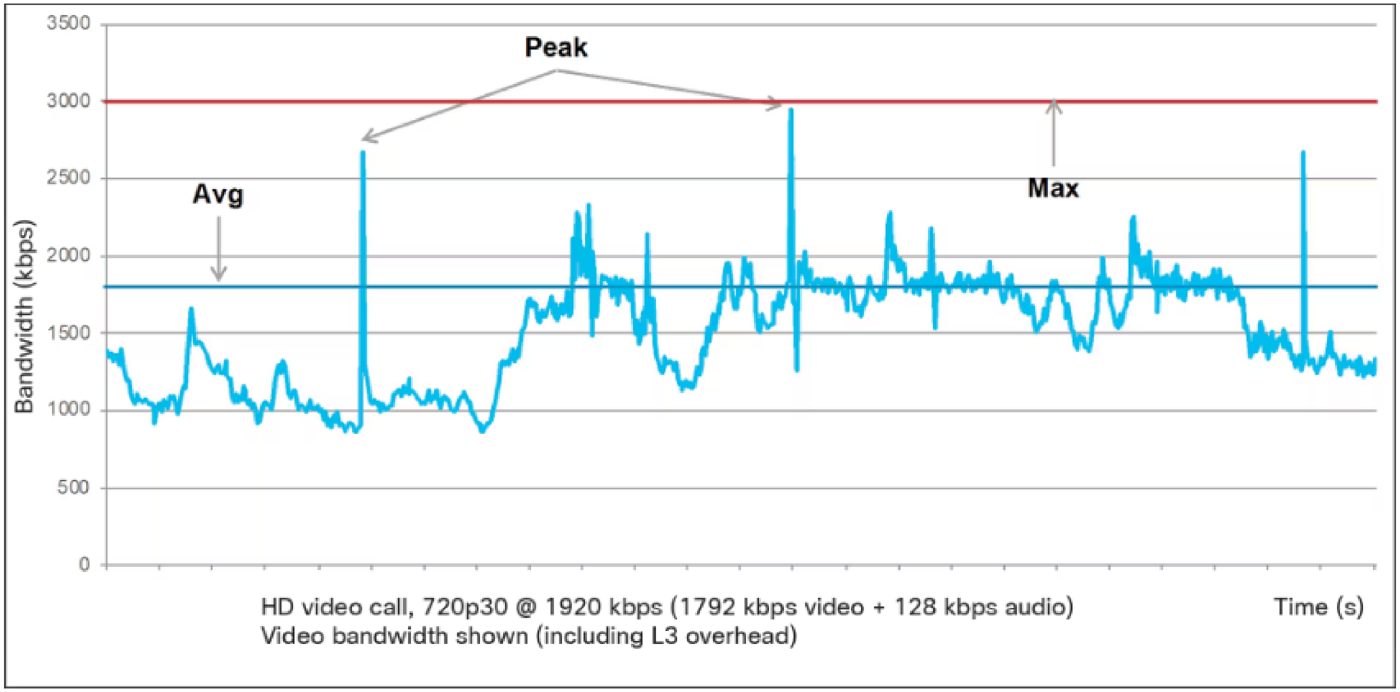
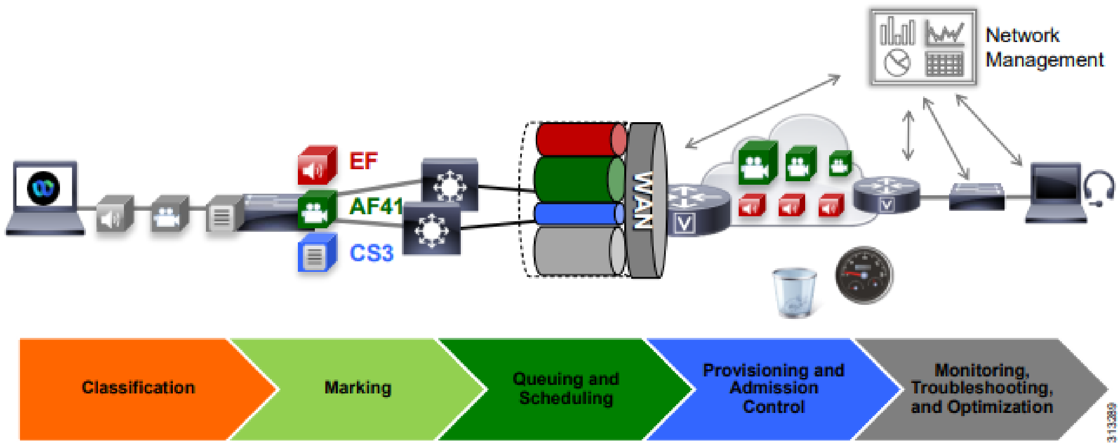


Figure 1.

Video Traffic: Bandwidth Usage High-definition Video Call

(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

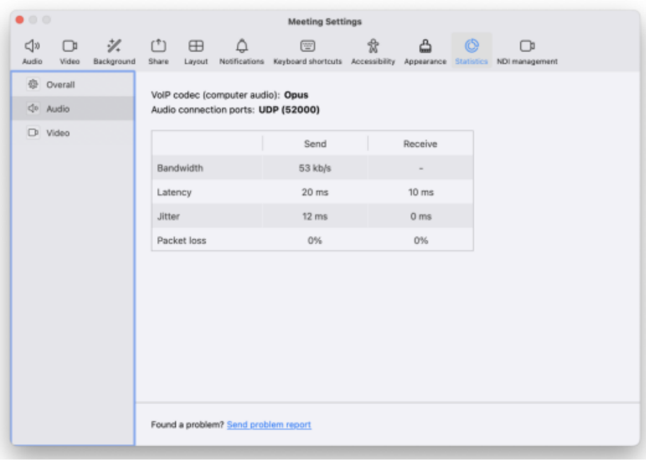
Figure 1 Architecture for Bandwidth Management

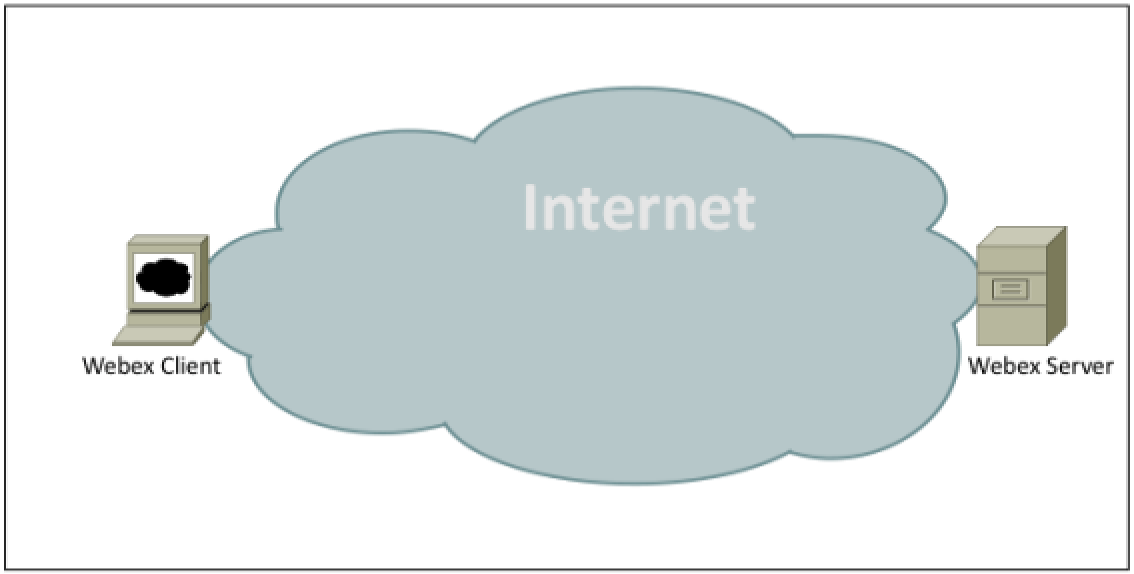


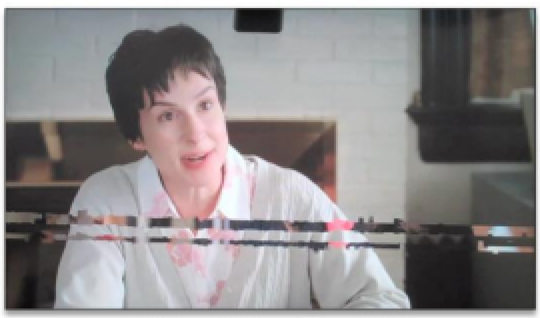
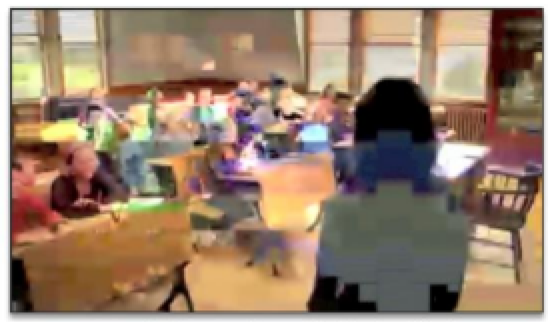
Recommended Deployment

Modify the existing on-premises QoS switch and WAN and Internet policies to include Webex Services identification, classification, and marking.

- Identify and classify media and signaling traffic from the Webex App and associated workloads as well as Webex Devices.
- Media and signaling marking recommendations:
  1. Mark all audio with Expedited Forwarding class EF (includes all audio of voice-only calls as well as audio for all types of video calls).
  2. Mark all Webex App video with an Assured Forwarding class of AF41 for a prioritized video class of service and AF42 for an opportunistic class of service. The marking of AF41 or AF42 will depend on the choice of whether to deploy opportunistic video during the on-premises deployment phase.
  3. Mark all call signaling with CS3. (All call signaling in HTTPS traffic will be marked based on the enterprise's current policy of traffic marking for HTTP/HTTPS – unless using NBAR which is covered in detail below)
- Configure QoS on all media originating and terminating applications such as the Video Mesh Nodes, Expressway and Cisco Unified Border Element.
- Update the WAN edge ingress re-marking policy.
- Update the WAN edge egress queuing and scheduling policy.

	Every Cisco video endpoint employs several smart media techniques to avoid network congestion, recover from packet loss, and optimize network resources. The following smart media techniques are just some of the techniques employed by Cisco Webex App and Devices: Media resilience techniques - Encoder pacing - Forward Error Correction (FEC) - Rate adaptation (https://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Collaboration/AltDesigns/BWM-Wbx.pdf.) **Poor Audio / Video Quality** - Full-featured Meetings Help > Health Checker > Audio and Video Statistics... • Indicates TCP or UDP w/ Source Port • Latency / Packet Loss / Jitter  (https://www.ciscolive.com/c/dam/r/ciscolive/global-event/docs/2024/pdf/BRKCOL-3431.pdf.)

	There are three main factors that impact the quality of an audio or video call. These factors are **packet loss**, **latency**, and **jitter**. As shown in Figure 2, **packet loss is simply losing one or more packets within a stream of packets**. In this example, a Voice over IP (VoIP) packet is lost going from the Webex client to the Webex server. Loss can occur for a number of reasons in a network but most commonly it is caused by congestion or resource contention in the network or on the endpoints. This leads to routers, switches, or the endpoints themselves dropping or delaying packets. In other cases, packets encounter such a high delay that they are received too late to be played out. These packets will also be counted as lost.  ***Figure 2:*** *Packet Loss Example*

	***Note**: Control Hub Diagnostics highlights potential problems with packet loss and delay by coloring voice and video streams as Good, Fair, or Poor. It is important to understand that just because Control Hub Diagnostics flags part of a conversation as Fair or Poor, it does not necessarily mean that the user had a bad experience. Voice and video compensation mechanisms and algorithms in combination with other factors can mitigate the impact of loss and delay to the point where it is not noticeable by the user.* **Video Quality Artifacts** The first step in troubleshooting video quality is to clearly understand the symptom experienced by the end user. Getting a screenshot or a picture of the display screen usually gives more information than a verbal description of the problem. The type of video artifacts observed in a screenshot or picture can be used to determine the troubleshooting path. For example, in Figure 13, pictures (a) and (b) are artifacts that are caused when video streams experience packet loss. Just by seeing the picture, you can immediately start troubleshooting to identify the source of the packet loss.  a. Video with line stripe artifacts  b. Frozen video with block artifacts

Webex Control Hub makes it quite easy to find the device type and operating system version used by the end user to join a Webex meeting. All you need is the end user's email address and meeting time to find the right meeting and to get end user's device details. Figure 14 shows the list of meetings attended by a participant with the email address, rtpmsuser1@gmail.com.

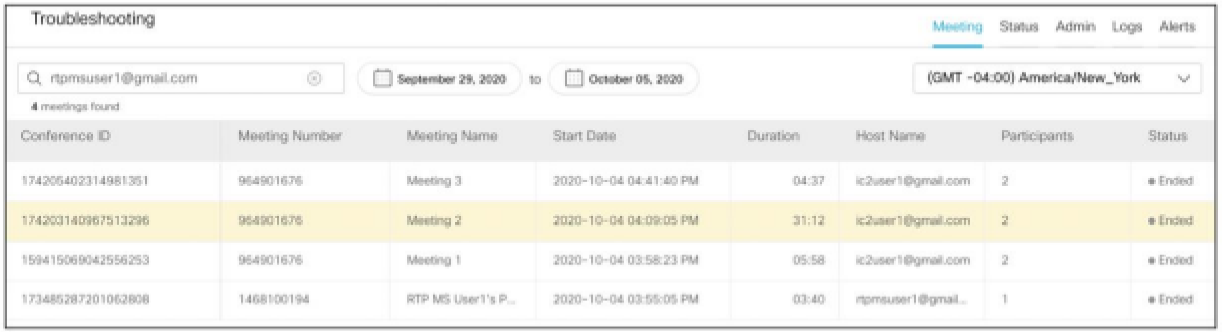


***Figure 14**: Participant Device Details*

(https://www.cisco.com/c/dam/td-xml/en_us/collaboration/cloud_cmr/pcia_2_0/reports/Troubleshooting_Audio_and_Video_Quality_Using_Webex_Control_Hub.pdf.)

forward the base layer from the video	The Accused Instrumentalities forward the base layer from the video router to at least two of the set of video receivers. For example, because the base layer is the minimum bitstream subset of a given scalable bitstream that is capable of being decoded, every receiver that receives a decodable bitstream subset is sent at least the base layer.

router to at least two of the set of video receivers;	**G.8.8.2 Specification of the base layer bitstream** Each scalable bitstream that conforms to this specification shall contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A. This base layer bitstream is derived by invoking the sub-bitstream extraction process as specified in clause G.8.8.1 with dIdTarget being equal to 0 and qIdTarget being equal to 0 and the base layer bitstream being the output. NOTE – Although all scalable bitstreams that conform to one or more of the profiles specified in this annex contain a base layer bitstream that conforms to one or more of the profiles specified in Annex A, the complete scalable bitstream (prior to operation of the base layer extraction process specified in this clause) may not conform to any profile specified in Annex A. **Rec. ITU-T H.264 (06/2019)** 595
wherein the scalable video coding router forwards all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with	The Accused Instrumentalities forward all of the set of enhancement layers to at least two of the video receivers in the set of video receivers with bandwidth-sufficient conditions. For example, to all of the end destinations (video receivers) that can be reached at data speeds sufficient to transmit the complete scalable bitstream (i.e. can be reached with sufficient bandwidth), the scalable video router will select and send all of the enhancement layers of the scalable video bitstream. For instance, as shown below, all of the set of enhancement layers are forwarded when priority_id is less than or equal to pIdTarget, temporal_id is less than or equal to tIdTarget, and dependency_id is less than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is less than or equal to qIdTarget) for all VCL NAL units and filler data NAL units in the bitstream.

bandwidth-sufficient conditions,	The sub-bitstream is derived by applying the following operations in sequential order: 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G. See ITU-T H.264 at 489.

The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation.

See ITU-T H.264 at 624.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

Table 4. Webex Meetings Bandwidth per Resolution Table

Layer	Bandwidth Range
90p active thumbnail (each)	~60-100 kb/s
180p main video	125-200 kb/s
360p main video	470-640 kb/s
720p main video	900k-1.5 mb/s
Content sharing (sharpness, 1080p/5)	120k – 1.3 mb/s
Content sharing (motion, 720p/30)	900k – 2.5 mb/s

Webex Meetings Desktop App Bandwidth Controls

Webex administrators have 2 key controls to help control bandwidth as used by clients that connect to Webex meetings should they choose to. Namely, you can cap the meeting layouts at either 360p as the max available resolution, or to enable 720p layers. Whether your site is administered on Webex Control Hub or Webex Site Administrator, the following controls are available in Configuration > Common Site Settings > Options:

- [ ] Turn on high-quality video (360p) *(Meetings, Training, Events and Support)*
- [ ] Turn on high-definition video (720p) *(Meetings, Training and Events)*

Figure 5.

Webex Meetings Desktop App Bandwidth Controls

Webex Media Improvements

The following are media improvements that have occurred in releases from 40.7 – 40.10

Improved Video Experience over WIFI (Defer video down-speeding): WIFI environments present a specific set of challenges and through testing we noticed the algorithms were too quick to drop resolution/bandwidth in certain WIFI environments. In 40.7 the algorithms have been updated to 'defer the down-speeding" of video for brief WIFI specific network dropouts. This resulted in improved user experience over WIFI. See figure5 for an illustration of this. When latency spikes for a specific duration the improved algorithms can detect that this is a WIFI specific event and not thus delay the down-speeding of the video resolution until that duration (n) is exceeded in which case would infer a different network event increasing latency.

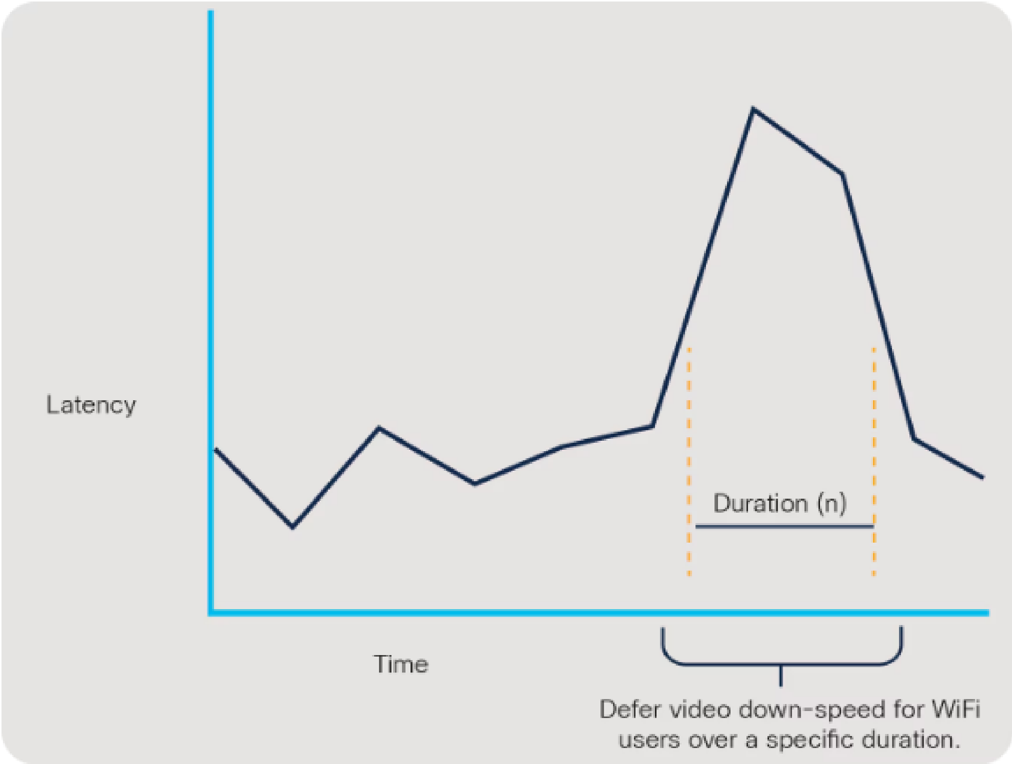


Figure 6.

Deferred Video Down-speeding

Video Super Scaling is a technique used for users on sub-optimal networks. We provide 720p like experiences to users who are experiencing network impairments resulting in 360p connections to our servers for their meeting. Our updated scaling algorithm provides enhanced video quality when this degradation occurs, making network impairment downgrades less quality impacting. If a 360p image is simply scaled geometrically to a 720p resolution, the quality has not improved, but our technique involves migrating video post-processing tasks to the client for some video flows versus doing it in the cloud, resulting in being able to keep more precious bits of information to reconstruct the video at a given bandwidth.

In a meeting as shown in figure 8 and 9, the bandwidth requirements are as follows. Assuming the receiver is not bandwidth restricted, Webex app will adapt to available network conditions. Table 8 outlines the bandwidth utilization. Figure 13 illustrates the Active Speaker layout which is the default meeting layout for Webex Teams App.

Table 8. Cisco Webex Teams Bandwidth Utilization

Meeting Scenario	Main Video+Audio (no content sharing)	Main Video+Audio (content sharing)
Webex Teams 1:1 Call	1.5 mb/s peak, 1.3 mb/s avg	1.5 mb/s peak, 1.3 mb/s avg
Webex Teams Meeting (fig 13 layout), 6 participants	2.3 mb/s peak, 1.7 mb/s avg	2.9/ mb/s peak, 2.0 mb/s avg
Webex Teams Meeting (fig 14 layout), 6 participants	2.5 mb/s peak, 1.8 mb/s avg	3.5 mb/s peak, 2.3 mb/s avg

	(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)
wherein the scalable video coding router selectively forwards one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the remaining video receivers in the set of video receivers based upon the identified bandwidth-	The Accused Instrumentalities selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the remaining video receivers in the set of video receivers based upon the identified bandwidth-limited conditions. For example, if a given end destination can be reached at data speeds insufficient to transmit the complete scalable bitstream because of a bandwidth limitation, the scalable video router will select and send all of the enhancement layers which can be sent at sufficient data speeds for each of the plurality of video receivers. For instance, as shown below, fewer than all of the set of enhancement layers are forwarded when priority_id is greater than pIdTarget, or temporal_id is greater than tIdTarget, or dependency_id is greater than dIdTarget (or dependency_id is equal to dIdTarget but quality_id is greater than qIdTarget) for any VCL NAL units or filler data NAL units in the bitstream.

limited conditions;	The sub-bitstream is derived by applying the following operations in sequential order: 1. Mark all VCL NAL units and filler data NAL units for which any of the following conditions are true as "to be removed from the bitstream": – priority_id is greater than pIdTarget, – temporal_id is greater than tIdTarget, – dependency_id is greater than dIdTarget, – dependency_id is equal to dIdTarget and quality_id is greater than qIdTarget. 2. Remove all access units for which all VCL NAL units are marked as "to be removed from the bitstream". 3. Remove all VCL NAL units and filler data NAL units that are marked as "to be removed from the bitstream". 4. When dIdTarget is equal to 0 and qIdTarget is equal to 0, remove the following NAL units: – all NAL units with nal_unit_type equal to 14 or 15, – all NAL units with nal_unit_type equal to 6 in which the first SEI message has payloadType in the range of 24 to 35, inclusive. 5. Remove all NAL units with nal_unit_type equal to 6 that only contain SEI messages that are part of a scalable nesting SEI message with any of the following properties: – sei_temporal_id is greater than tIdTarget, – the minimum value of (sei_dependency_id[i] << 4) + sei_quality_id[i] for all i in the range of 0 to num_layer_representations_minus1, inclusive, is greater than (dIdTarget << 4) + qIdTarget. 6. Remove all NAL units with nal_unit_type equal to 6 that contain SEI messages with payloadType equal to 24, 28, or 29. See ITU-T H.264 at 595. All sub-bitstreams that can be derived using the sub-bitstream extraction process as specified in clause G.8.8.1 with any combination of values for priority_id, temporal_id, dependency_id, or quality_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Annexes A and G.

See ITU-T H.264 at 595.

The representation of a particular scalable layer is the set of NAL units that represents the set union of the particular scalable layer and all scalable layers on which the particular scalable layer directly or indirectly depends. The representation of a scalable layer is also referred to as scalable layer representation. In the following specification of this clause, the terms representation of a scalable layer and scalable layer representation are also used for referring to the access unit set that can be constructed from the NAL units of the scalable layer representation. A scalable layer representation can be decoded independently of all NAL units that do not belong to the scalable layer representation. The decoding result of a scalable layer representation is the set of decoded pictures that are obtained by decoding the access unit set of the scalable layer representation.

See ITU-T H.264 at 595.

sciencedirect.com/topics/engineering/scalable-coding

to serve the various users' requirements. Usually all participants would have to agree on some minimal quality for the video conference, but this is not really satisfactory in many cases. With the advent of SVC, the MCU is not needed because transcoding is not needed. Using the SVC standard, so-called *video routers* can be substituted for the MCUs. These routers only have to forward and delete packets as appropriate to that user's connection and needs, so little extra delay is added. In effect, video packets can be routed by the video router based on their headers, just like other network packets.

	For example, the Accused Instrumentalities selectively forward one or more of the set of enhancement layers, but fewer than all of the set of enhancement layers, to at least two of the remaining video receivers in the set of video receivers based upon the identified bandwidth-limited conditions:

Table 4. Webex Meetings Bandwidth per Resolution Table

Layer	Bandwidth Range
90p active thumbnail (each)	~60-100 kb/s
180p main video	125-200 kb/s
360p main video	470-640 kb/s
720p main video	900k-1.5 mb/s
Content sharing (sharpness, 1080p/5)	120k – 1.3 mb/s
Content sharing (motion, 720p/30)	900k – 2.5 mb/s

Webex Meetings Desktop App Bandwidth Controls

Webex administrators have 2 key controls to help control bandwidth as used by clients that connect to Webex meetings should they choose to. Namely, you can cap the meeting layouts at either 360p as the max available resolution, or to enable 720p layers. Whether your site is administered on Webex Control Hub or Webex Site Administrator, the following controls are available in Configuration > Common Site Settings > Options:

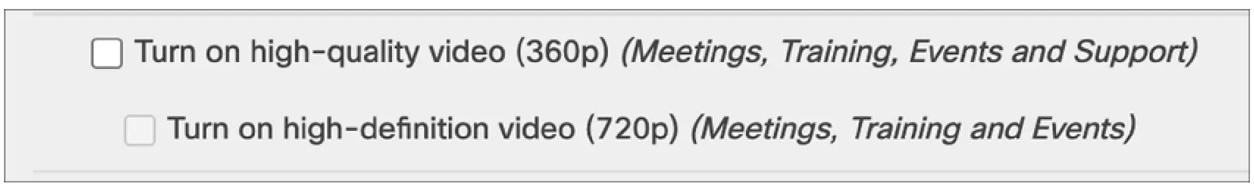


Figure 5.

Webex Meetings Desktop App Bandwidth Controls

Webex Media Improvements

The following are media improvements that have occurred in releases from 40.7 – 40.10

Improved Video Experience over WIFI (Defer video down-speeding): WIFI environments present a specific set of challenges and through testing we noticed the algorithms were too quick to drop resolution/bandwidth in certain WIFI environments. In 40.7 the algorithms have been updated to 'defer the down-speeding" of video for brief WIFI specific network dropouts. This resulted in improved user experience over WIFI. See figure5 for an illustration of this. When latency spikes for a specific duration the improved algorithms can detect that this is a WIFI specific event and not thus delay the down-speeding of the video resolution until that duration (n) is exceeded in which case would infer a different network event increasing latency.

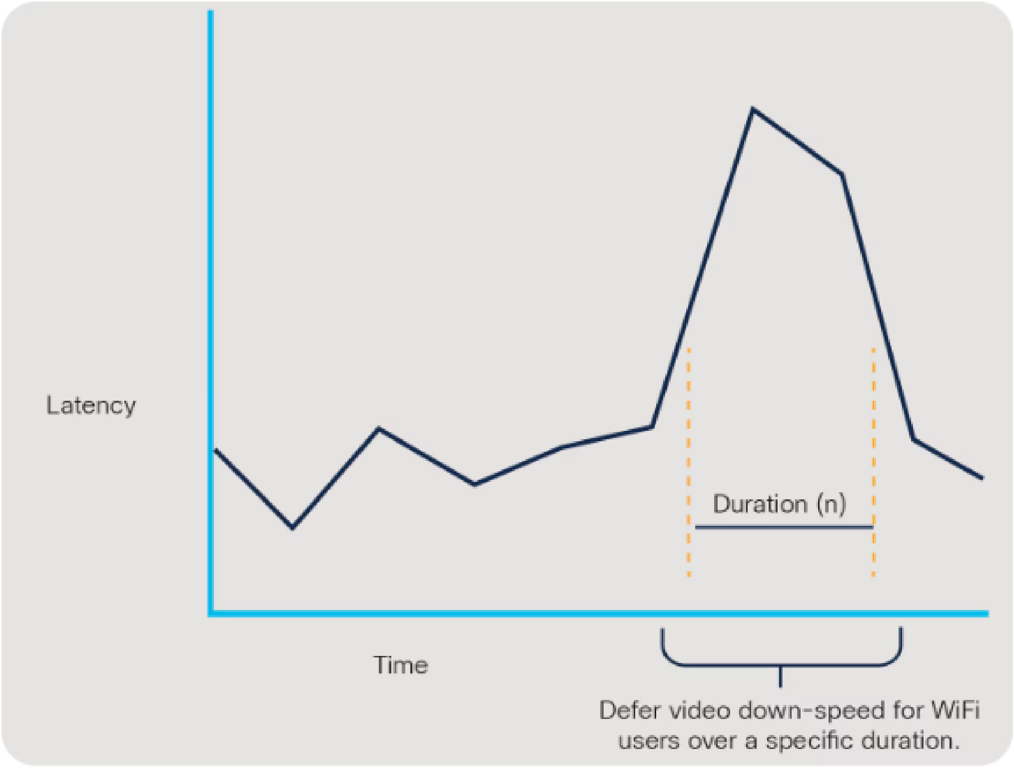


Figure 6.

Deferred Video Down-speeding

Video Super Scaling is a technique used for users on sub-optimal networks. We provide 720p like experiences to users who are experiencing network impairments resulting in 360p connections to our servers for their meeting. Our updated scaling algorithm provides enhanced video quality when this degradation occurs, making network impairment downgrades less quality impacting. If a 360p image is simply scaled geometrically to a 720p resolution, the quality has not improved, but our technique involves migrating video post-processing tasks to the client for some video flows versus doing it in the cloud, resulting in being able to keep more precious bits of information to reconstruct the video at a given bandwidth.

In a meeting as shown in figure 8 and 9, the bandwidth requirements are as follows. Assuming the receiver is not bandwidth restricted, Webex app will adapt to available network conditions. Table 8 outlines the bandwidth utilization. Figure 13 illustrates the Active Speaker layout which is the default meeting layout for Webex Teams App.

Table 8. Cisco Webex Teams Bandwidth Utilization

Meeting Scenario	Main Video+Audio (no content sharing)	Main Video+Audio (content sharing)
Webex Teams 1:1 Call	1.5 mb/s peak, 1.3 mb/s avg	1.5 mb/s peak, 1.3 mb/s avg
Webex Teams Meeting (fig 13 layout), 6 participants	2.3 mb/s peak, 1.7 mb/s avg	2.9/ mb/s peak, 2.0 mb/s avg
Webex Teams Meeting (fig 14 layout), 6 participants	2.5 mb/s peak, 1.8 mb/s avg	3.5 mb/s peak, 2.3 mb/s avg

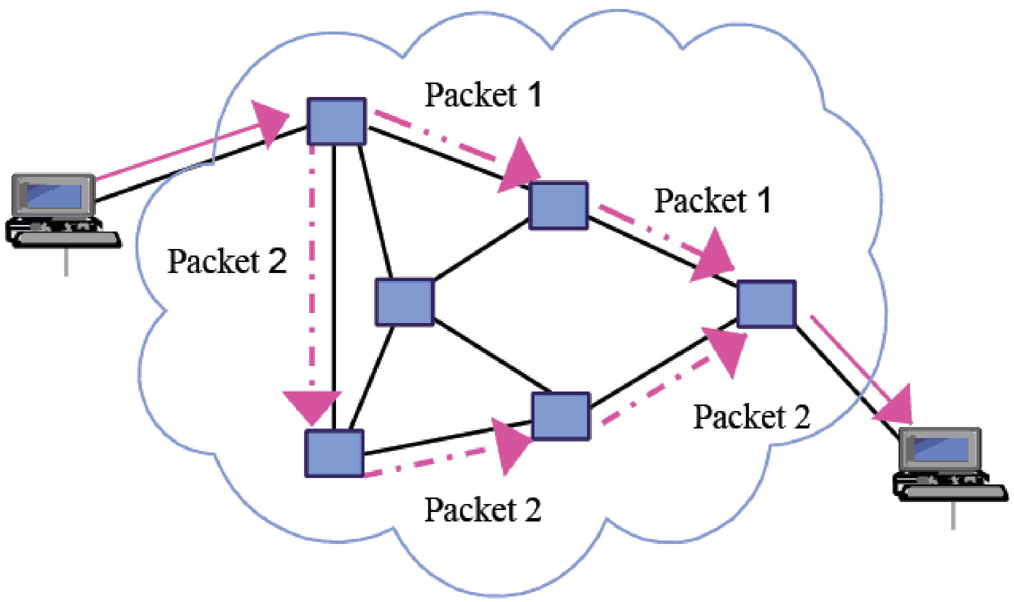
(https://www.cisco.com/c/en/us/products/collateral/conferencing/webex-meetings/white_paper_c11-691351.html.)

wherein the layered video data stream is transmitted according to an internet protocol, and wherein each layer of the layered video data stream comprises data packets,	The Accused Instrumentalities include a scalable video coding router, wherein the layered video data stream is transmitted according to an internet protocol, and wherein each layer of the layered video data stream comprises data packets. For example, scalable video routers facilitate the distribution of scalable bitstreams of video information across a packet switched network, such as the internet. Data packets are the result of a process of dividing information into units so that said units may be transmitted across a packet switched network. At the destination, the data packets are reassembled to produce the information from which said data packets were derived.

networkencyclopedia.com/packet-switching/

What is Packet Switching?

Packet Switching is the process by which a networking or telecommunications device accepts a packet and switches it to a telecommunications device that will take it closer to its destination. Packet switching allows data to be sent over the telecommunications network in short bursts or "packets" that contain sequence numbers so that they can be reassembled at the destination.



Data packets that are transmitted across the internet are comprised of a header and payload/body, according to the internet protocol. The payload is comprised of the data intended for transmission, and the header is information facilitating said transmission. Moreover, the payload of a packet is often another packet, comprising another header and another payload, as each packet layer is designed to be handled by a different part of the data distribution process. (NOTE: packet layers are not to be confused with scalable video layers).

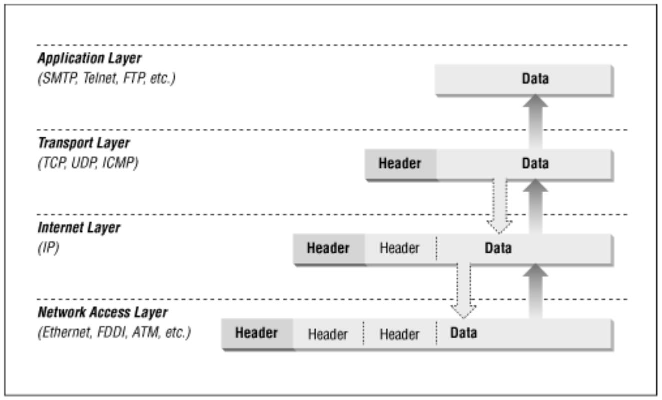
Not Secure | web.deu.edu.tr/doc/oreily/networking/firewall/ch06_03.htm

Packets are constructed in such a way that layers for each protocol used for a particular connection are wrapped around the packets, like the layers of skin on an onion.

At each layer, a packet has two parts: the header and the body. The header contains protocol information relevant to that layer, while the body contains the data for that layer which often consists of a whole packet from the next layer in the stack. Each layer treats the information it gets from the layer above it as data, and applies its own header to this data. At each layer, the packet contains all of the information passed from the higher layer; nothing is lost. This process of preserving the data while attaching a new header is known as *encapsulation*.

At the application layer, the packet consists simply of the data to be transferred (for example, part of a file being transferred during an FTP session). As it moves to the transport layer, the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP) preserves the data from the previous layer and attaches a header to it. At the next layer, IP considers the entire packet (consisting now of the TCP or UDP header and the data) to be data, and now attaches its own IP header. Finally, at the network access layer, Ethernet or another network protocol considers the entire IP packet passed to it to be data, and attaches its own header. Figure 6.2 shows how this works.

Figure 6.2: Data encapsulation

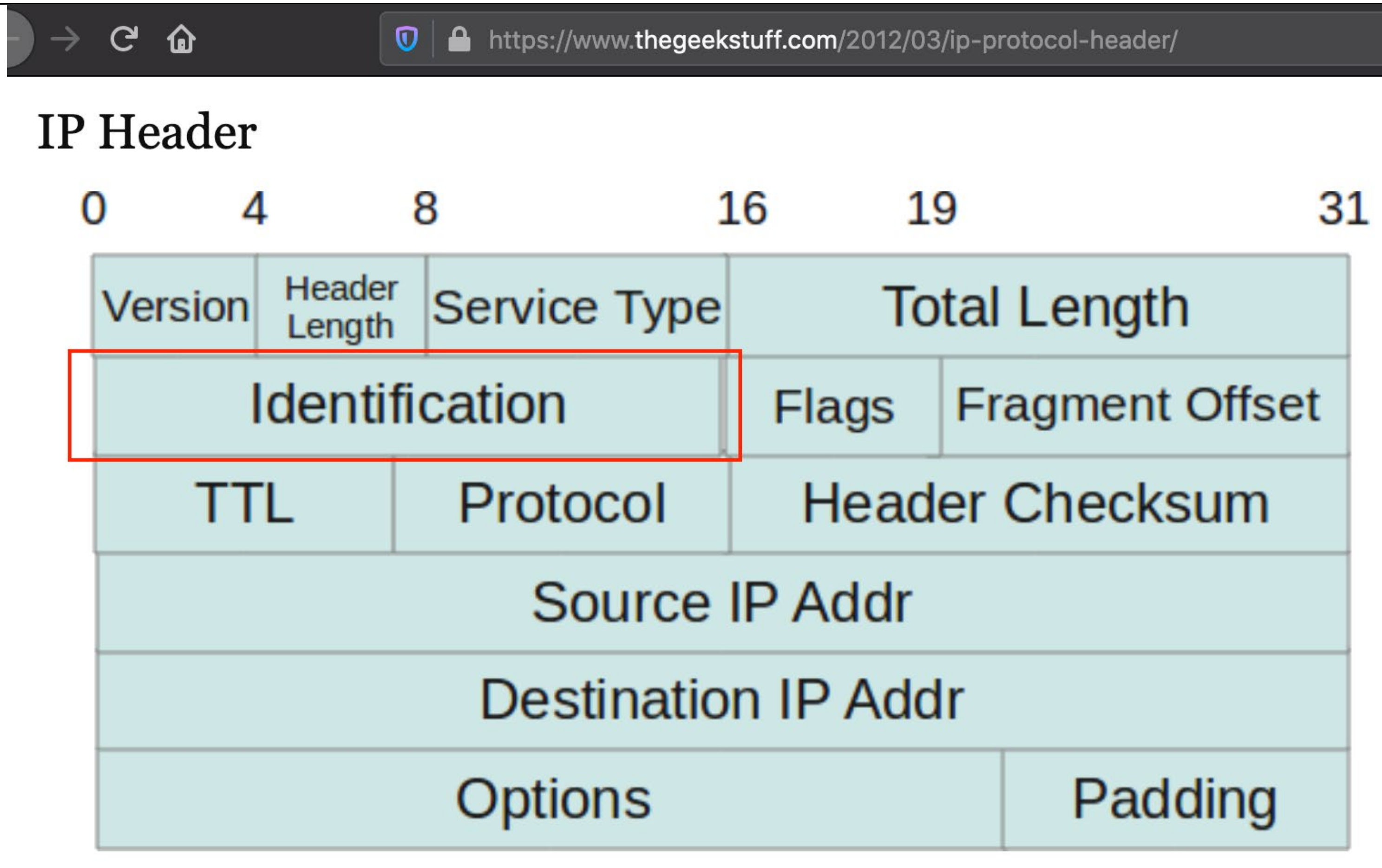


At the other side of the connection, this process is reversed. As the data is passed up from one layer to the next higher layer, each header (each skin of the onion) is stripped off by its respective layer. For example, the Internet layer removes the IP header before passing the encapsulated data up to the transport layer (TCP or UDP).

Additionally, the final payload of a scalable video data packet is a NAL Unit, which is also a unit of information comprising a header and a body. Scalable bitstreams that conform to the H.264 standard are transmitted as a sequence of NAL units.

	ip.hhi.de/imagecom_G1/savce/downloads/SVC-Overview.pdf *A. Network Abstraction Layer (NAL)* The coded video data are organized into NAL units, which are packets that each contains an integer number of bytes. A NAL unit starts with a one-byte header, which signals the type of the contained data. The remaining bytes represent payload data. NAL units are classified into VCL NAL units, which con-
each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a	The Accused Instrumentalities include a scalable video coding router, wherein each layer of the layered video data stream comprises data packets, each of which is encoded with a sequence number and a layer identifier, and wherein the layer identifier for each data packet is based upon a layer to which the packet belongs. For example, each data packet is identified by a sequence number which uniquely identified said packet among all other packets in the message. Every packet in the sequence contains an identification value that is one more than that of the previous packet in the sequence.

layer to which the packet belongs	ⓘ Not Secure | linfo.org/packet_header.html **Packet Header Definition** A *packet header* is the portion of an IP (Internet protocol) *packet* that precedes its body and contains addressing and other data that is required for it to reach its intended destination. Packets are the fundamental unit of information transport in all modern computer networks, and increasingly in other communications networks as well. They can be a fixed size or variable sizes, depending on the system. Regardless of their size, each packet consists of three main parts: a header, the body, also called the *payload*, and a *trailer*. The header's format is specified in the Internet protocol. It normally contains 20 bytes of data, although an option exists within it that allows the addition of more bytes. Among the contents of the header are the version of IP (which is always set to 4, because IPv4 is being used), the sender's IP address, the intended receiver's IP address, the number of packets the message has been broken into, the identification number of the particular packet, the protocol (e.g. 1 for ICMP, 2 for IGMP, 6 for TCP and 17 for UDP) used, the packet length (on networks that have variable length packets), the *time to live* (i.e., the number of links or *hops* that the packet can be routed before being allowed to expire) and synchronization data (several bits that help the packet match up to the network).


https://www.thegeekstuff.com/2012/03/ip-protocol-header/
IP Header
0
4
8
16
19
31
Version
Header Length
Service Type
Total Length
Identification
Flags
Fragment Offset
TTL
Protocol
Header Checksum
Source IP Addr
Destination IP Addr
Options
Padding

https://www.thegeekstuff.com/2012/03/ip-protocol-header/

- ***Identification(16 bits)***: This field is used for uniquely identifying the IP datagrams. This value is incremented every-time an IP datagram is sent from source to the destination. This field comes in handy while reassembly of fragmented IP data grams.

For example, each data packet is encoded with a layer identifier, wherein the layer identifier for each data packet is based upon a layer to which the packet belongs.

Each scalable layer is associated with a unique layer identifier as specified later in this clause. The representation of a particular scalable layer with a particular layer identifier layerId does not include any scalable layer with a layer identifier greater than layerId, but it may include scalable layers with layer identifiers less than layerId. The scalable layers on which a particular scalable layer depends may be indicated in the scalability information SEI message as specified later in this clause.

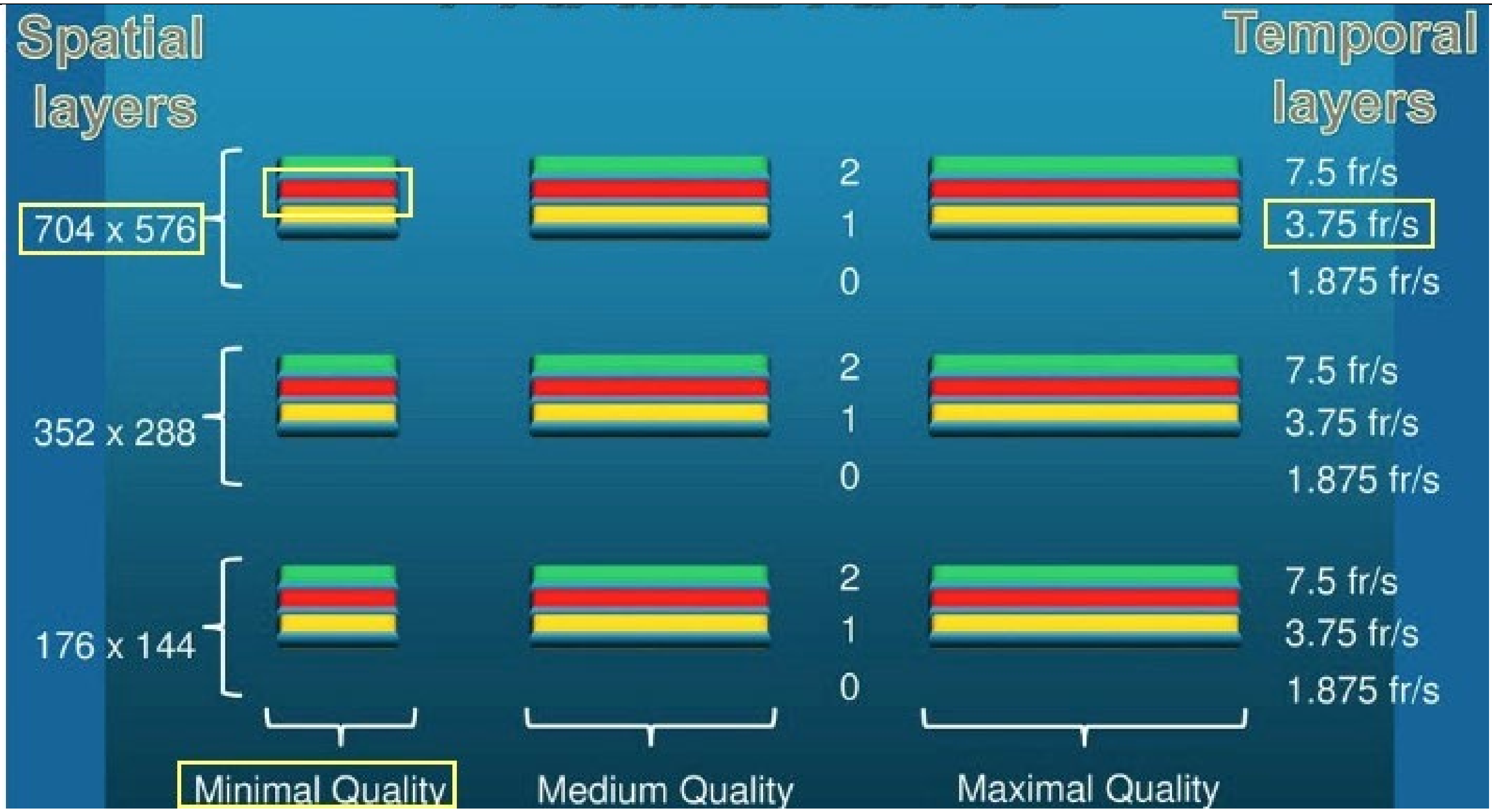
NOTE 3 – When all scalable layers for which scalability information is provided in the scalability information SEI message have sub_pic_layer_flag[i] equal to 0, the unique layer identifier values may be set equal to (128 * dependency_id + 8 * quality_id + temporal_id), with dependency_id, quality_id, and temporal_id being the corresponding syntax elements that are associated with the VCL NAL units of a scalable layer.

See ITU-T H.264 at 624.

NAL unit headers for scalable bitstreams that comply with the H.264 standard comprise a layer identifier, which is comprised of the set of values including priority_id dependency_id, quality_id, and temporal_id, according to the standard.

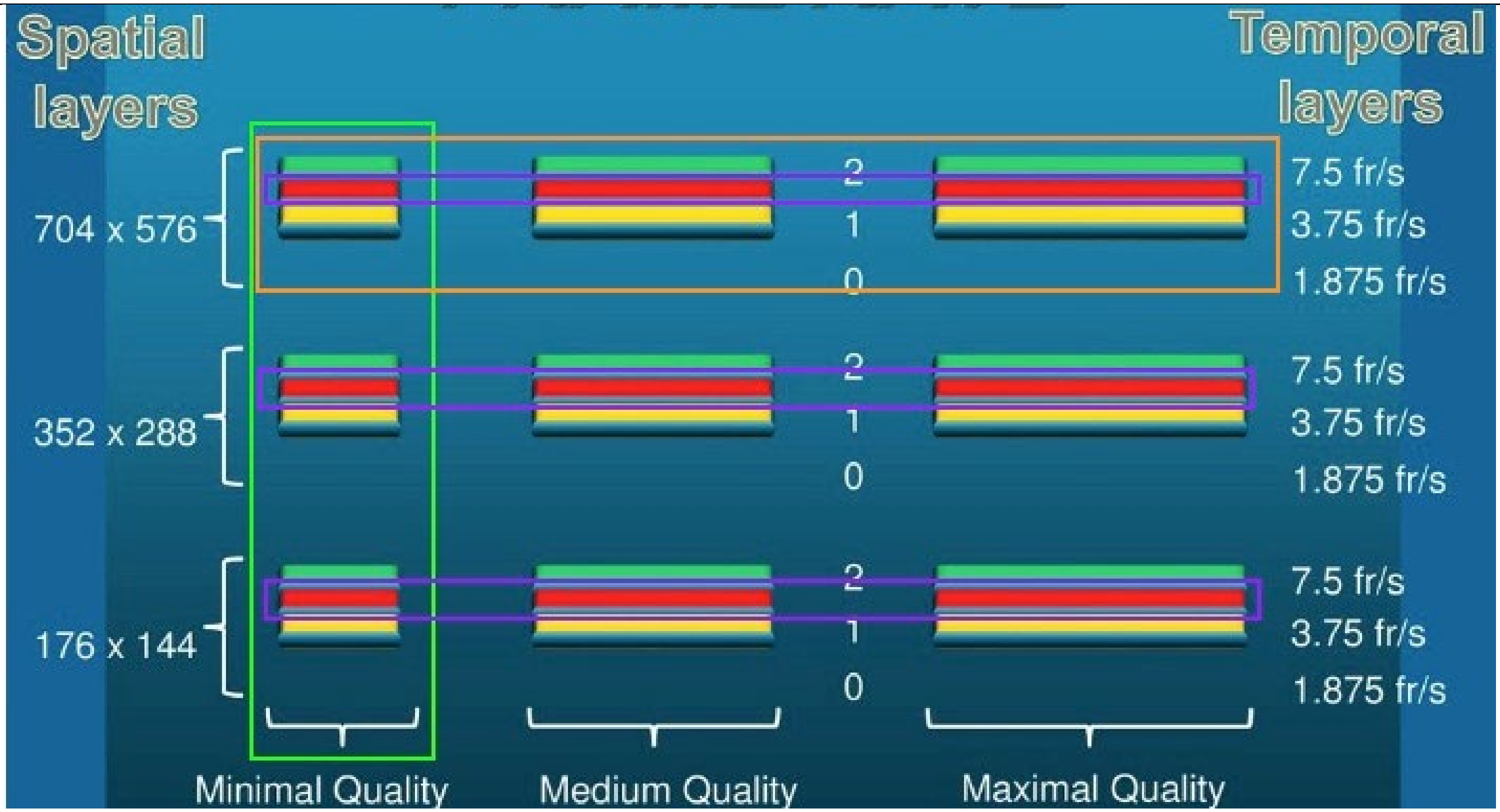
dependency_id specifies a dependency identifier for the NAL unit. dependency_id shall be equal to 0 in prefix NAL units. The assignment of values to dependency_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1.

	See ITU-T H.264 at 459. **priority_id** specifies a priority identifier for the NAL unit. The assignment of values to priority_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. See ITU-T H.264 at 458. **quality_id** specifies a quality identifier for the NAL unit. quality_id shall be equal to 0 in prefix NAL units. The assignment of values to quality_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The variable DQId is derived by DQId = (dependency_id << 4) + quality_id (G-63) When nal_unit_type is equal to 20, the bitstream shall not contain data that result in DQId equal to 0. See ITU-T H.264 at 459. **temporal_id** specifies a temporal identifier for the NAL unit. The assignment of values to temporal_id is constrained by the sub-bitstream extraction process as specified in clause G.8.8.1. The value of temporal_id shall be the same for all prefix NAL units and coded slice in scalable extension NAL units of an access unit. When an access unit contains any NAL unit with nal_unit_type equal to 5 or idr_flag equal to 1, temporal_id shall be equal to 0. See ITU-T H.264 at 459. For instance, the graphic below illustrates an example of a hierarchal virtual layer structure. The layers scale regarding multiple elements, such as resolution (spatial), frame rate (temporal), quality, etc. Every unique combination of the values that comprise the layer identifier indicates the relative location of a virtual layer within the hierarchy. For instance, the in the illustration below the convergence of the spatial value for “704 x 576”, the quality value for “Minimal Quality”, and the temporal value for “3.75 fr/s” comprises a single layer within the hierarchy.



Source: https://slideplayer.com/slide/13389021/.

For instance, in the example below there are nine layers that are possible for each spatial value (e.g. 704 x 576), or for each quality value (e.g. minimal quality), or for each temporal value (e.g. 3.75 fr/s). Bitstream subsets can be derived from all of the NAL units of a given layer and all layers with a layer identifier less than the layer identifier of said layer, forming a hierarchy of virtual layers.



Source, https://slideplayer.com/slide/13389021/.